International Journal of Approximate Reasoning 50 (2009) 979–999

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

# Monotonic Variable Consistency Rough Set Approaches

Jerzy Błaszczyński [a], Salvatore Greco [b], Roman Słowiński [a,c,*], Marcin Szeląg [a]

[a] Institute of Computing Science, Poznań University of Technology, 60-965 Poznań, Poland
[b] Faculty of Economics, University of Catania, Corso Italia, 55, 95129 Catania, Italy
[c] Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland

ARTICLE INFO

*Article history:*
Received 11 July 2008
Received in revised form 9 January 2009
Accepted 25 February 2009
Available online 14 March 2009

*Keywords:*
Rough sets
Dominance-based Rough Set Approach
Monotonicity
Bayesian confirmation
Rough membership
Likelihood
Accuracy of approximation
Variable precision
Variable consistency

ABSTRACT

We consider probabilistic rough set approaches based on different versions of the definition of rough approximation of a set. In these versions, consistency measures are used to control assignment of objects to lower and upper approximations. Inspired by some basic properties of rough sets, we find it reasonable to require from these measures several properties of monotonicity. We consider three types of monotonicity properties: monotonicity with respect to the set of attributes, monotonicity with respect to the set of objects, and monotonicity with respect to the dominance relation. We show that consistency measures used so far in the definition of rough approximation lack some of these monotonicity properties. This observation led us to propose new measures within two kinds of rough set approaches: Variable Consistency Indiscernibility-based Rough Set Approaches (VC-IRSA) and Variable Consistency Dominance-based Rough Set Approaches (VC-DRSA). We investigate properties of these approaches and compare them to previously proposed Variable Precision Rough Set (VPRS) model, Rough Bayesian (RB) model, and previous versions of VC-DRSA.



## 1. Introduction

Calculation of rough set approximations can be the first step of the analysis of data. It allows to identify consistent data and put them into lower approximations of sets (concepts, classes, or unions of ordered classes). The following step is usually related to generalization of data. When we consider a classification problem, this step consists in induction of a classifier that can be further used for prediction. In the original rough set approach proposed by Pawlak [19,20], and in the dominance-based rough set approach proposed by Greco et al. [8,9,11,24], the lower approximation of a set is defined by a strict inclusion relation of some granules of knowledge in the approximated set. The lower approximation is thus composed of the granules that are subsets of the approximated set. Other granules are not included into lower approximation, regardless of the size of their overlap with the set and/or its complement. This definition of the lower approximation appears to be too restrictive in practical applications. In consequence, lower approximations of sets are often empty, preventing generalization of data in terms of relative certainty. This observation has motivated research on probabilistic generalizations of rough sets.

Different versions of probabilistic rough set approaches were proposed, starting from Variable Precision Rough Set (VPRS) model [26,28,29], Variable Consistency Dominance-based Rough Set Approaches (VC-DRSA) [1,9,10], Bayesian Rough Set model and Rough Bayesian (RB) model [25,26], decision theoretic rough set model [13,30,31] and Parameterized Rough Sets [14]. The probabilistic rough set approaches allow to extend lower approximation of a set by objects with sufficient evidence

* Corresponding author. Address: Institute of Computing Science, Poznań University of Technology, 60-965 Poznań, Poland.
*E-mail addresses:* jblaszczynski@cs.put.poznan.pl (J. Błaszczyński), salgreco@unict.it (S. Greco), rslowinski@cs.put.poznan.pl (R. Słowiński), mszelag@cs.put.poznan.pl (M. Szeląg).

0888-613X/$ - see front matter 
doi:10.1016/j.ijar.2009.02.011

for membership to the set. To quantify this evidence, the authors propose different measures of the overlap between a granule of knowledge based on a considered object and the approximated set or its complement. We call such measures *consistency measures*.

Inspired by some basic properties of rough sets, we find it reasonable to require from consistency measures several properties of monotonicity that correspond directly to monotonicity properties of the lower approximation. The present paper focuses on three types of monotonicity properties. These monotonicity properties are considered in various dimensions of the analyzed data set and are related to:

(1) extension of the set of attributes,
(2) extension of the set of objects,
(3) extension of the union of ordered classes,
(4) improvement of evaluation of an object.

Monotonicity in dimension (1) requires that precisiation of the description of objects by addition of attributes can only give more evidence for the assignment of these objects to the approximated set. Precisiation means here a more detailed description of objects, without considering a semantic value of the additional information. Let us observe that if a semantic value of additional attributes would be considered, then the precisiation could decrease the evidence for the assignment of objects to the approximated set. For example, a semantic value of additional attributes could depend on whether the precisiation by these attributes decreases or increases the confusion related to the assignment of objects to the approximated set. Then, in the first case, the semantic value would be considered positive, and in the second, negative. Thus, additional attributes with a negative semantic value would not increase the evidence for the assignment of objects to the approximated set. Monotonicity in dimension (1) is concordant with monotonicity of the accuracy of approximation defined by Pawlak [20]. This type of monotonicity is desirable for reduction of attributes in probabilistic rough set approaches. From classification perspective, monotonicity in this dimension corresponds to reasonable outcome of classifiers that are induced from data sets that overlap in the dimension of attributes (i.e., multiple classifiers generated on overlapping subsets of attributes from the extended set of attributes).

Monotonicity in dimension (2) requires that extension of the approximated set by addition of new objects, should not negatively affect the evidence for membership of the "old" objects to the approximated set. From classification perspective, this property allows to generate compatible classifiers on overlapping subsets of objects from the extended set of objects (i.e., incremental classifiers or ensembles of classifiers created on overlapping sets of objects, like in bagging or boosting). Let us observe that monotonicity in dimension (2) may be discussed in the context of Bayesian confirmation theory, i.e., the theory which studies how a piece of evidence $E$ provides "evidence for or against" or "support for or against" hypothesis $H$ (for an extensive survey see [5]). In fact, we can imagine that new objects constitute new evidence which may confirm or disconfirm the hypothesis that an object can be assigned to an approximated set. Let us explain this point in terms of the famous paradox, called black raven paradox [15]. Consider hypothesis $H \equiv$ "all ravens are black", which corresponds to the idea of assigning to set "black" all objects having value "raven" on attribute "raven yes or not". Hypothesis $H$ can be read as the implication "if an object is a raven, than it is black". Within the rough set approach (for a discussion about relationships between Bayesian confirmation theory and rough set theory see [12]), the hypothesis concerns the membership to decision class "black" of those objects which according to the condition attribute are "raven". At a first look, one can imagine that any object being "raven" and "black" confirms the hypothesis, any object being "raven" and "non-black" disconfirms the hypothesis, and all objects being not "raven" do not confirm and do not disconfirm the hypothesis. In general, this observation can be expressed as follows: an object confirms an implication if and only if it satisfies both the premise and the conclusion of the implication (i.e., it is "black" and "raven"); it disconfirms the implication if and only if it satisfies the premise, but not the conclusion (i.e., it is "non-black" but it is "raven"); it does not confirm and does not disconfirm the implication if it does not satisfy the premise (i.e., it is not "raven"). In this perspective, each new black object cannot disconfirm the hypothesis. In fact, it can confirm the hypothesis if it is also "raven" or neither confirm nor disconfirm if it is not "raven". In our context, this means that extending the approximated set of black objects, we cannot reduce the membership of an object to the considered set, and this agrees with monotonicity in dimension (2). Hempel observed in [15] that hypothesis $H$ is logically equivalent to the implication "if an object is non-black, then it is not raven", which is confirmed by objects being "non-black" and not "raven". Remark that this observation leads to the paradox that pink socks can confirm the hypothesis that "all ravens are black". Also in this case, a black object cannot disconfirm the hypothesis (even if it cannot also confirm it), because it does not satisfy the premise. In our context, this agrees again with monotonicity in dimension (2). Observe, however, that in case of probabilistic confirmation, some authors find it reasonable to expect that "black non-ravens" can reduce the confirmation degree [12]. In this case, a considered confirmation measure of the hypothesis "if $\Phi$, then $\Psi$", can be expressed as credibility of the proposition "if $\Psi$ is satisfied more frequently when $\Phi$ is satisfied rather than when $\Phi$ is not satisfied". According to this understanding, black ravens and non-black non-ravens confirm the hypothesis, while non-black ravens and black non-ravens disconfirm the hypothesis. With respect to black non-ravens, they disconfirm the hypothesis because they increase the probability that $\Psi$ is satisfied when $\Phi$ is not satisfied, i.e., they increase the probability that an object is black when it is not raven. In this sense, expectations for probabilistic confirmation do not agree with monotonicity in dimension (2).

Finally, monotonicity with respect to the dominance relation is considered in dimensions (3) and (4). Monotonicity in these dimensions concerns data sets with specified orders of preference. They allow to generate classifiers that permit to make classification decisions respecting preference orders. Property (m3) is related to an important property of DRSA, which wants that a lower approximation of any upward (downward) union of ordered classes includes a lower approximation of any of its upward (downward) sub-unions. Property (m4) ensures that if an object belongs to a lower approximation of an upward (downward) union of ordered classes, then all objects from this union which dominate (are dominated by) this object will also belong to the lower approximation.

At the end of this introduction, it is worth noting, however, that instead of requiring monotonicity properties from consistency measures, one could accept non-monotonic behavior of consistency measures and consider application of non-monotonic logic [3,7,18]. This way of looking at probabilistic generalizations of rough sets could be an interesting subject for future research.

In the next section, we remind basic definitions of original Indiscernibility-based Rough Set Approach and Dominance-based Rough Set Approach. Then, we define monotonicity properties required for consistency measures that are used in Monotonic Variable Consistency Rough Set Approaches. In Section 3, we show which of the monotonicity properties are satisfied by consistency measures that were used in probabilistic rough set approaches proposed so far. We also give examples of shortcomings of these measures. In Section 4, we define new monotonic Variable Consistency Indiscernibility-based Rough Set Approaches (VC-IRSA). Along the way, two types of monotonic consistency measures are introduced. We prove and interpret their properties. In Section 5, we show how the measures defined for the indiscernibility relation can be reformulated for the dominance relation. In consequence, new monotonic Variable Consistency Dominance-based Rough Set Approaches (VC-DRSA) are proposed. Finally, we present an illustrative example for indiscernibility-based and dominance-based approaches. We conclude by giving remarks and recommendations for applications of the new approaches.

## 2. Monotonicity properties required for rough set approaches

In the rough set approach, classification of object $y$ from universe $U$ to a given set $X \subseteq U$ is based on available data. Data is presented as a decision table, where rows correspond to objects from $U$ and columns correspond to attributes from a finite set $A$. Among attributes from set $A$ there are attributes with preference-ordered value sets, called criteria, and regular attributes whose value sets are not preference-ordered. Moreover, the set of attributes $A$ is divided into disjoint sets of condition attributes $C$ and decision attributes $D$. For simplicity, we assume set $D$ to be a singleton $D = \{d\}$.

The decision attribute $d$ makes a partition of set $U$ into a finite number of disjoint sets of objects, called decision classes. Let $X \subseteq U$ be one of these decision classes. Decision about classification of object $y \in U$ to set $X$ depends on its class label known from the decision table, and/or on its relation with other objects from the table. In the original rough set approach, the considered relation is the *indiscernibility relation* [19,20]. For this reason, we call this approach Indiscernibility-based Rough Set Approach (IRSA). Consideration of the indiscernibility relation is meaningful when set of attributes $A$ is composed of regular attributes only. Indiscernibility relation makes a partition of universe $U$ into disjoint blocks of objects that have the same description and are considered indiscernible. Such blocks are called *granules*. Let $V_{a_i}$ be the value set of attribute $a_i \in C$ and $f : U \times C \to V_{a_i}$ be a total function such that $f(x, a_i) \in V_{a_i}$. Indiscernibility relation $I_P$ is defined for a non-empty subset of attributes $P \subseteq C$ as

$$I_P = \{(y,z) \in U \times U : f(y,a_i) = f(z,a_i) \text{ for all } a_i \in P\}.$$

Moreover, $I_P(y)$ denotes a set of objects indiscernible with object $y$ using set of attributes $P$. It is called a granule of $P$-indiscernible objects.

When condition attributes from $C$ and decision attribute $d$ have preference-ordered value sets, in order to make meaningful classification decisions, one has to consider the *dominance relation* instead of the indiscernibility relation. It has been proposed in [8,9,11,24] and the resulting approach was called Dominance-based Rough Set Approach (DRSA). Dominance relation makes a partition of universe $U$ into granules being *dominance cones*. The dominance relation $D_P$ is defined for a non-empty subset of criteria $P \subseteq C$ as

$$D_P = \{(y,z) \in U \times U : f(y,a_i) \succeq f(z,a_i) \text{ for all } a_i \in P\},$$

where $f(y,a_i) \succeq f(z,a_i)$ means "$y$ is at least as good as $z$ with respect to (w.r.t.) criterion $a_i$". Dominance relation $D_P$ is a partial preorder (i.e. reflexive and transitive). For each object $y \in U$ two dominance cones (granules) are defined w.r.t. $P \subseteq C$. The $P$-positive dominance cone $D_P^+(y)$ is composed of all objects that are dominating $y$. The $P$-negative dominance cone $D_P^-(y)$ is composed of all objects that are dominated by $y$. Formal definitions of dominance cones are as follows:

$$D_p^+(y) = \{z \in U : zD_Py\},$$
$$D_p^-(y) = \{z \in U : yD_Pz\}.$$

We are considering a classification problem with $n$ disjoint classes. While in IRSA, decision classes $X_i$, $i = 1, \dots, n$, are not necessarily ordered, in DRSA, they are ordered, such that if $i < j$, then class $X_i$ is considered to be worse than $X_j$. Moreover, DRSA takes into account monotonic relationships between evaluations of objects on particular criteria and assignment of these objects into decision classes. For example, the better the value of criterion $a_i \in C$ for object $y$, the better the decision class it may belong. From this follows the *dominance principle* which says that if evaluations of object $y$ on all considered criteria are not worse than evaluations of object $z$, then $y$ should be assigned to a class not worse than $z$. Violation of this principle causes *inconsistency* in the data table which is captured within DRSA by approximations of sets. In order to handle preference orders, and monotonic relationships between evaluations on criteria and assignment to decision classes, approximations made in DRSA concern the following unions of decision classes: upward unions $X_i^{\geqslant} = \bigcup_{t \geqslant i} X_t$, where $i = 2, 3, \ldots, n$, and downward unions $X_i^{\leqslant} = \bigcup_{t \leqslant i} X_t$, where $i = 1, 2, \ldots, n-1$.

One of the most important features of rough set approaches is the separation of knowledge which is consistent, from knowledge which is possibly inconsistent. In IRSA, a key point is to find evidence for assignment of objects to particular decision classes $X_i$. In DRSA, the key point is to find evidence for assignment of objects to unions of decision classes $X_i^{\geqslant}$ and $X_i^{\leqslant}$.

In order to avoid repetition of the same definitions and properties for IRSA and DRSA, we will use a unique symbol $X$ to denote a set of all objects belonging to class $X_i$, in the context of IRSA, or to union of classes $X_i^{\geqslant}, X_i^{\leqslant}$, in the context of DRSA.

Let us specify conditions that must be satisfied by consistency measures. We distinguish gain-type and cost-type consistency measures. First, let us consider $y_1, y_2 \in U$, $P \subseteq C$, $X \subseteq U$. Given description of $y_1$ and $y_2$ by $P$:

- a gain-type consistency measure $f_X^P(y)$ is any measure satisfying condition: $f_X^P(y_1) \geqslant f_X^P(y_2) \iff$ it is not less likely that $y_1$ belongs to $X$, than that $y_2$ belongs to $X$,
- a cost-type consistency measure $g_X^P(y)$ is any measure satisfying condition: $g_X^P(y_1) \leqslant g_X^P(y_2) \iff$ it is not less likely that $y_1$ belongs to $X$, than that $y_2$ belongs to $X$.

Second, let us consider $y \in U$, $P \subseteq C$, $X, Y \subseteq U$, where $Y$ has the same interpretation as $X$ (i.e., it denotes a class or a union of classes). Given description of $y$ by $P$:

- a gain-type consistency measure $f_X^P(y)$ is any measure satisfying condition: $f_X^P(y) \geqslant f_Y^P(y) \iff$ it is not less likely that $y$ belongs to $X$, than that it belongs to $Y$.
- a cost-type consistency measure $g_X^P(y)$ is any measure satisfying condition: $g_X^P(y) \leqslant g_Y^P(y) \iff$ it is not less likely that $y$ belongs to $X$, than that it belongs to $Y$.

A consistency measure expresses the evidence for membership to set $X$. For a gain-type measure, the higher the value, the more consistent is the given object. For a cost-type measure, the lower the value, the more consistent is the given object. In this paper, we investigate desirable properties of consistency measures.

Each set $X$, may include objects for which, due to inconsistency, we are unable to find enough evidence for their membership to $X$. In such a case, we can approximate set $X$ by two sets, the *P-lower approximation* and the *P-upper approximation* of $X$, where $P \subseteq C$. Let us give generic definitions of $P$-lower approximations of set $X$, which involve *consistency measures* $f_X^P(y)$ or $g_X^P(y)$.

For $P \subseteq C, X \subseteq U, \ y \in U$, given a gain-type consistency measure $f_X^P(y)$ and a gain-threshold $\alpha_X$, we get the following definitions of $P$-lower approximation of set $X$:

$$\underline{P}^{\alpha_X}(X) = \{y \in U : f_X^P(y) \geqslant \alpha_X\} \tag{1}$$
$$\text{or } \underline{P}^{\alpha_X}(X) = \{y \in X : f_X^P(y) \geqslant \alpha_X\}. \tag{2}$$

Analogically, given a cost-type consistency measure $g_X^P(y)$ and a cost-threshold $\beta_X$, we get the following definitions:

$$\underline{P}^{\beta_X}(X) = \{y \in U : g_X^P(y) \leqslant \beta_X\} \tag{3}$$
$$\text{or } \underline{P}^{\beta_X}(X) = \{y \in X : g_X^P(y) \leqslant \beta_X\}. \tag{4}$$

In the above definitions, gain-threshold $\alpha_X \in [0, A_X]$ and cost-threshold $\beta_X \in [0, B_X]$. These thresholds are parameters depending on the interpretation of the gain-type or cost-type consistency measure, respectively. They play the role of technical parameters influencing the degree of consistency of objects belonging to lower approximation of $X$.

Thus, the values of $A_X$ and $B_X$ also depend on the interpretation of the corresponding consistency measure. For example, in case of probabilistic $P$-lower approximation defined using the rough membership measure, $A_X = 1$ and value of gain-threshold $\alpha_X \in [0, 1]$ can be calculated using method presented in [13,30]. This method is based on application of the Bayesian decision procedure in transformation of risk into the value of $\alpha_X$.

The above definitions of $P$-lower approximations relax the non-parametric definitions. Precisely, the non-parametric definition for IRSA and class $X_i$ is as follows:

$$\underline{P}(X_i) = \{y \in U : I_P(y) \subseteq X_i\} = \{y \in X_i : I_P(y) \subseteq X_i\},$$

and for DRSA, and unions of classes $X_i^{\geqslant}, X_i^{\leqslant}$, it is as follows:

$$\underline{P}(X_i^{\geqslant}) = \{y \in U : D_P^+(y) \subseteq X_i^{\geqslant}\} = \{y \in X_i^{\geqslant} : D_P^+(y) \subseteq X_i^{\geqslant}\},$$
$$\underline{P}(X_i^{\leqslant}) = \{y \in U : D_P^-(y) \subseteq X_i^{\leqslant}\} = \{y \in X_i^{\leqslant} : D_P^-(y) \subseteq X_i^{\leqslant}\}.$$

An obvious condition of this relaxation is:

$$\underline{P}(X) \subseteq \underline{P}^{\alpha_X}(X), \tag{5}$$
$$\underline{P}(X) \subseteq \underline{P}^{\beta_X}(X). \tag{6}$$

The definition of $P$-upper approximation and the definition of $P$-boundary of set $X$ make use of the complementarity property of rough approximations, and are the same for all the approaches considered in this work.

For $P \subseteq C, X, \neg X \subseteq U$, where $\neg X = U - X$, $P$-upper approximation of set $X$ is defined as

$$\overline{P}^{\alpha_X}(X) = U - \underline{P}^{\alpha_X}(\neg X), \quad \overline{P}^{\beta_X}(X) = U - \underline{P}^{\beta_X}(\neg X), \tag{7}$$

while $P$-boundary of set $X$ is defined as

$$Bn_P^{\alpha_X}(X) = \overline{P}^{\alpha_X}(X) - \underline{P}^{\alpha_X}(X), \quad Bn_P^{\beta_X}(X) = \overline{P}^{\beta_X}(X) - \underline{P}^{\beta_X}(X). \tag{8}$$

Let us remark that the notion of consistency was also used in IRSA, to measure consistency of the whole decision table [4,16,22,23]. In this case, different instances of the entropy measure were applied instead of the quality of approximation. Entropy measures were also applied to define consistency of a granule composed of $P$-indiscernible objects [23]. In the case of the whole decision table, as well as in the case of a single granule, consistency was considered with respect to all possible classes from the decision table.

In the present paper, we understand consistency in a different way. We consider consistency of particular objects with respect to the approximated sets.

One can observe that properties of rough approximations defined above depend on properties of consistency measures $f_X^P(y)$ and $g_X^P(y)$. Thus, it is possible to formulate some properties with respect to these measures, which ensure desirable properties of rough approximations.

For IRSA and DRSA, it is reasonable to require that consistency measures $f_X^P(y)$ and $g_X^P(y)$ fulfill the following properties of monotonicity (henceforth called *monotonicity properties*):

(m1) Monotonicity w.r.t. set of attributes $P \subseteq C$. Formally, for all $P \subseteq P' \subseteq C, X \subseteq U, y \in U$, a gain-type measure $f_X^P(y)$ is monotonically non-decreasing w.r.t. $P$, if and only if (iff)

$$f_X^P(y) \leqslant f_X^{P'}(y), \tag{9}$$

and a cost-type measure $g_X^P(y)$ is monotonically non-increasing w.r.t. $P$, iff

$$g_X^P(y) \geqslant g_X^{P'}(y). \tag{10}$$

(m2) Monotonicity w.r.t. set of objects $X \subseteq U$, when set $X$ is augmented by new objects. Formally, for all $P \subseteq C$, $X \subseteq U$, $X' = X \cup X^{\Delta}$, $X^{\Delta} \cap U = \emptyset$, $y \in U$, a gain-type measure $f_X^P(y)$ is monotonically non-decreasing w.r.t. $X$, iff

$$f_X^P(y) \leqslant f_{X'}^P(y), \tag{11}$$

and a cost-type measure $g_X^P(y)$ is monotonically non-increasing w.r.t. $X$, iff

$$g_X^P(y) \geqslant g_{X'}^P(y). \tag{12}$$

Moreover, for DRSA, it is reasonable to require that measures $f_{X_i^{\geqslant}}^P(y)$ (or $f_{X_i^{\leqslant}}^P(y)$) and $g_{X_i^{\geqslant}}^P(y)$ (or $g_{X_i^{\leqslant}}^P(y)$) fulfill the following monotonicity properties:

(m3) Monotonicity w.r.t. union of classes $X_i^{\geqslant} \subseteq U$ and $X_k^{\leqslant} \subseteq U$. Formally, for all $P \subseteq C$, $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, $j \leqslant i$, $X_k^{\leqslant} \subseteq X_l^{\leqslant} \subseteq U$, $l \geqslant k$, $y \in U$, gain-type measures $f_{X_i^{\geqslant}}^P(y)$ and $f_{X_k^{\leqslant}}^P(y)$ are monotonically non-decreasing w.r.t. $X_i^{\geqslant}$ and $X_k^{\leqslant}$, respectively, iff

$$f_{X_i^{\geqslant}}^P(y) \leqslant f_{X_j^{\geqslant}}^P(y), \quad f_{X_k^{\leqslant}}^P(y) \leqslant f_{X_l^{\leqslant}}^P(y). \tag{13}$$

Analogously, a cost-type measures $g_{X_i^{\geqslant}}^P(y)$ and $g_{X_k^{\leqslant}}^P(y)$ are monotonically non-increasing w.r.t. $X_i^{\geqslant}$ and $X_k^{\leqslant}$, respectively, iff

$$g_{X_i^{\geqslant}}^P(y) \geqslant g_{X_j^{\geqslant}}^P(y), \quad g_{X_k^{\leqslant}}^P(y) \geqslant g_{X_l^{\leqslant}}^P(y). \tag{14}$$

(m4) Monotonicity w.r.t. $P$-dominance relation, $P \subseteq C$. Formally, for all $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$, and $*$ standing for either $\geqslant$ or $\leqslant$ in every instance, a gain-type measure $f_{X_i^*}^P(y)$ is monotonically non-decreasing w.r.t. $P$-dominance relation, iff

$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow f_{X_i^*}^P(y_1) \geqslant f_{X_i^*}^P(y_2), \tag{15}$$

and a cost-type measure $g_{X_i^*}^P(y)$ is monotonically non-increasing w.r.t. $P$-dominance relation, iff

$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow g_{X_i^*}^P(y_1) \leqslant g_{X_i^*}^P(y_2). \tag{16}$$

Monotonicity properties (m1) and (m2) are related to the basic properties of rough sets. Monotonicity properties (m3) and (m4) are specific to DRSA. A rough set approach is called monotonic when the consistency measure used to define its lower approximation fulfills relevant monotonicity properties. For IRSA, relevant properties are (m1) and (m2), while for DRSA, relevant properties are (m1), (m2), (m3) and (m4).

Property (m1) is particularly important. Property (m1) of measures $f_X^P(y)$ and $g_X^P(y)$ ensures monotonicity of $P$-lower approximation w.r.t. set of attributes $P \subseteq C$, defined according to (2) and (4), respectively. This property imposes that additional information about objects from $U$ can only give more evidence for the observed assignment of objects to classes. In this case, additional information means a precisiation by more detailed description of considered objects using an extended set of attributes. Property (m1) is also concordant with the observation that additional attributes can only decrease comparability in the set of objects. When less objects are comparable, then also less inconsistent assignments to classes is observed.

Property (m2) of measures $f^P_X(y)$ and $g^P_X(y)$ ensures monotonicity of $P$-lower approximation w.r.t. set of objects $X \subseteq U$. Property (m2) states that when we consider two sets of objects $X' \supset X$, the evidence for membership to $X'$ for objects from $X$ should not be worse than the evidence for their membership to $X$. In other words, extension of class $X_i$ or union of classes $X_i^{\geqslant}$ ($X_i^{\leqslant}$) by addition of new objects, should not negatively affect the evidence for membership of the objects to the extended class or union of classes.

In DRSA, property (m3) of measures $f^P_{X_i^{\geqslant}}(y)$ (or $f^P_{X_i^{\leqslant}}(y)$) and $g^P_{X_i^{\geqslant}}(y)$ (or $g^P_{X_i^{\leqslant}}(y)$) ensures monotonicity of $P$-lower approximation w.r.t. union $X_i^{\geqslant} \subseteq U$ (or $X_i^{\leqslant} \subseteq U$). This property states that value of a gain-type consistency measure for a union that is a superset should not decrease, while value of a cost-type consistency measure should not increase. For example, for object $y$ which belongs to upward unions $X_i^{\geqslant}$ and $X_j^{\geqslant}$, where $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, value of gain-type consistency measure $f^P_{X_j^{\geqslant}}(y)$ should not be worse than the value of this measure calculated for union $X_i^{\geqslant}$.

The importance of property (m4) in Variable Consistency DRSA (VC-DRSA) was already discussed in [1], however, under the name of *monotonicity of membership to lower approximation*. Monotonicity w.r.t. $P$-dominance relation, $P \subseteq C$, is a very desirable property for a measure used in the definition of $P$-lower approximation of union $X_i^*$, where $^*$ stands for either $\geqslant$ or $\leqslant$. In case of definitions based on formula (2), where it is checked if $f^P_{X_i^*}(y) \geqslant \alpha_{X_i^*}$, a consistency measure defined for $X_i^{\geqslant}$ should satisfy (15), while a consistency measure defined for $X_i^{\leqslant}$ should satisfy (16). For definitions based on formula (4), where it is checked if $g^P_{X_i^*}(y) \leqslant \beta_{X_i^*}$, a consistency measure defined for $X_i^{\geqslant}$ should satisfy (16), while a consistency measure defined for $X_i^{\leqslant}$ should satisfy (15). This ensures a kind of continuity of lower approximations – as soon as some object $y \in X_i^{\geqslant}$ is included in the $P$-lower approximation of union $X_i^{\geqslant}$, every object $z \in X_i^{\geqslant}$, which $P$-dominates $y$, will also be included in this approximation. Analogically, if some object $y \in X_i^{\leqslant}$ is included in $P$-lower approximation of union $X_i^{\leqslant}$, then every object $z \in X_i^{\leqslant}$, which is $P$-dominated by $y$, will also belong to the considered approximation.

## 3. Are rough membership, confirmation measures and Bayes factor monotonic consistency measures?

Rough membership measure was introduced in [27] and its properties were further investigated in [21,31]. It is used to control positive regions in Variable Precision Rough Set (VPRS) model [26,28,29] and in previous versions of Variable Consistency Dominance-based Rough Set Approaches (VC-DRSA) [1,9,10]. Rough membership was also considered in the context of attribute reduction [17].

In IRSA, rough membership of $y \in U$ to $X \subseteq U$ w.r.t. $P \subseteq C$ is defined as

$$\mu^P_X(y) = \frac{|I_P(y) \cap X|}{|I_P(y)|}.$$

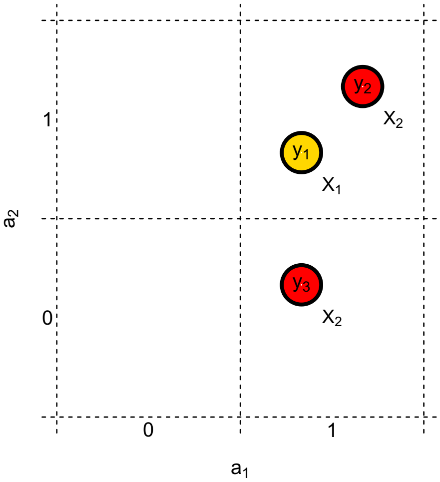
Rough membership is a gain-type consistency measure. It captures a ratio of objects that belong to granule $I_P(y)$ and to considered set $X$, among all objects belonging to granule $I_P(y)$. For example, if we would consider a medical diagnosis, the value of rough membership would express the ratio of the number of patients that have the same symptoms and suffer from the considered disease to the number of all patients that have the same symptoms. This measure can also be treated as an estimate of conditional probability $Pr(x \in X | x \in I_P(y))$. In IRSA, rough membership is used in definition (1), and it is expected to have properties (m1) and (m2). Unfortunately, property (m1) does not hold, which is shown by the example presented in Fig. 1. First, we consider attribute $a_1$ only. All objects have the same value on that attribute (i.e., they all belong to the same granule). Thus, $\mu^{\{a_1\}}_{X_2}(y_1) = \mu^{\{a_1\}}_{X_2}(y_2) = \mu^{\{a_1\}}_{X_2}(y_3) = 0.66$. Second, we consider set $P = \{a_1, a_2\}$. Then, we have two granules. The first one consists of objects $y_1, y_2$ and the other one is composed of object $y_3$. The value of rough membership to class $X_2$ drops to 0.5 in the first granule. On the other hand, property (m2) holds for rough membership measure $\mu^P_X(y)$ (see Proof of Theorem 4.8 in the Appendix).

Other measures than rough membership have also been used in rough set approaches. For example, confirmation measures [5,12] were considered together with rough membership in Parameterized Rough Sets (PRS) [14]. Confirmation measures quantify the degree to which membership of object $y$ to given granule $I_P(y)$ provides "evidence for or against" or "support for or against" assignment to considered set $X$. They are gain-type consistency measures and according to [14], they are used within definition (1). Confirmation measures should have properties (m1) and (m2). Unfortunately, as it may be shown, the well-known confirmation measures do not have property (m1).

The Bayes factor has similar properties to confirmation measures (its formulation is close to the confirmation measure $l$ [5]). It is a gain-type consistency measure used in the Rough Bayesian (RB) model [25]. The Bayes factor for $y \in U$ and $X \subseteq U$, w.r.t. $P \subseteq C$, is defined as

$$B^P_X(y) = \frac{|I_P(y) \cap X||\neg X|}{|I_P(y) \cap \neg X||X|}.$$

The Bayes factor is a ratio of estimates of two conditional probabilities $Pr(x \in I_P(y) | x \in X)$ and $Pr(x \in I_P(y) | x \in \neg X)$. Coming back to the example with medical diagnosis, the Bayes factor would express, in this case, the ratio of the estimate of probability that a patient has the considered symptoms on condition that he suffers from the considered disease to the estimate




| object | $a_1$ | $a_2$ | decision class |
|---|---|---|---|
| $y_1$ | 1 | 1 | $X_1$ |
| $y_2$ | 1 | 1 | $X_2$ |
| $y_3$ | 1 | 0 | $X_2$ |

**Fig. 1.** Exemplary set of objects described by set $P$ of attributes $a_1$ and $a_2$.

of probability that he has these symptoms on condition that he does not suffer from this disease. This measure is used in definition (1) and it is expected to have properties (m1) and (m2). Unfortunately, this is not the case. Let us come back to the example presented in Fig. 1. First, let us observe that $B^{\{a_1\}}_{X_2}(y_2) = 1$, while $B^P_{X_2}(y_2) = \frac{1}{2}$. This shows that the Bayes factor does not have property (m1). Second, let us extend the set of objects with one new object $y_4$, which belongs to class $X_2$ and has the following description: $a_1 = 0$, $a_2 = 1$. We can notice that $B^P_{X_2}(y_2) = \frac{1}{2}$ and $B^P_{X'_2}(y_2) = \frac{1}{3}$, where $X'_2 = \{y_2, y_3, y_4\}$. This shows that the Bayes factor also does not have property (m2).

Now, let us consider DRSA. In this case, rough membership is defined for $P \subseteq C$, $X^{\geqslant}, X^{\leqslant} \subseteq U$, $y \in U$, as

$$\mu^P_{X^{\geqslant}}(y) = \frac{|D^+_P(y) \cap X^{\geqslant}|}{|D^+_P(y)|}, \quad \mu^P_{X^{\leqslant}}(y) = \frac{|D^-_P(y) \cap X^{\leqslant}|}{|D^-_P(y)|},$$

where $X^{\geqslant}$, $X^{\leqslant}$ denote upward and downward unions of decision classes, respectively. Values of rough membership $\mu^P_{X^{\geqslant}}(y)$ and $\mu^P_{X^{\leqslant}}(y)$ can be interpreted as estimates of probability $Pr(z \in X^{\geqslant} | zD_P y)$ and $Pr(z \in X^{\leqslant} | yD_P z)$, respectively.

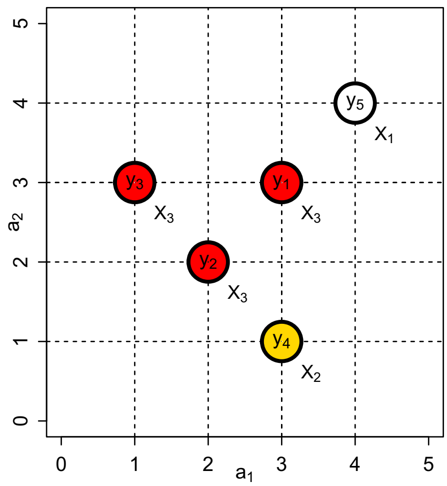
Formulation of the Bayes factor for $P \subseteq C$, $X^{\geqslant}, X^{\leqslant} \subseteq U$, $y \in U$, is as follows:

$$B^P_{X^{\geqslant}}(y) = \frac{|D^+_P(y) \cap X^{\geqslant}||\neg X^{\geqslant}|}{|D^+_P(y) \cap \neg X^{\geqslant}||X^{\geqslant}|}, \quad B^P_{X^{\leqslant}}(y) = \frac{|D^-_P(y) \cap X^{\leqslant}||\neg X^{\leqslant}|}{|D^-_P(y) \cap \neg X^{\leqslant}||X^{\leqslant}|}.$$

Both these measures are gain-type and they are used within DRSA in definition (2). They are expected to have properties (m1), (m2), (m3) and (m4). Measure $\mu^P_{X^{\geqslant}}(y)$ (or $\mu^P_{X^{\leqslant}}(y)$) has property (m2) – see Proof of Theorem 5.20 (or 5.21) in the Appendix. It also can be shown that measure $\mu^P_{X^{\geqslant}}(y)$ (or $\mu^P_{X^{\leqslant}}(y)$) has property (m3). Unfortunately, measure $\mu^P_{X^{\geqslant}}(y)$ (or $\mu^P_{X^{\leqslant}}(y)$) has neither property (m1) nor (m4). Moreover, measure $B^P_{X^{\geqslant}}(y)$ (or $B^P_{X^{\leqslant}}(y)$) has none of the monotonicity properties considered in this paper. Let us illustrate the lack of monotonicity by the example shown in Fig. 2. First, let us consider measure $\mu^P_{X^{\geqslant}}(y)$. We can notice that $\mu^{\{a_2\}}_{X^{\geqslant}_3}(y_2) = \frac{3}{4}$, while $\mu^P_{X^{\geqslant}_3}(y_2) = \frac{2}{3}$. Since $\mu^{\{a_2\}}_{X^{\geqslant}_3}(y_2) > \mu^P_{X^{\geqslant}_3}(y_2)$, measure $\mu^P_{X^{\geqslant}}(y)$ does not have property (m1). Moreover, $\mu^P_{X^{\geqslant}_3}(y_1) = \frac{1}{2}$ and $\mu^P_{X^{\geqslant}_3}(y_2) = \frac{2}{3}$, which shows that measure $\mu^P_{X^{\geqslant}}(y)$ also does not have property (m4). Second, let us consider measure $B^P_{X^{\geqslant}}(y)$. We can notice that $B^{\{a_2\}}_{X^{\geqslant}_3}(y_2) = 2$, while $B^P_{X^{\geqslant}_3}(y_2) = \frac{4}{3}$. Since $B^{\{a_2\}}_{X^{\geqslant}_3}(y_2) > B^P_{X^{\geqslant}_3}(y_2)$, measure $B^P_{X^{\geqslant}}(y)$ does not have property (m1). In order to show that measure $B^P_{X^{\geqslant}}(y)$ does not have property (m2), let us assume that object $y_3$ is not originally present in the considered data set and is added as a new object. We can observe that $B^P_{X^{\geqslant}_3}(y_2) = 2 > B^P_{X'^{\geqslant}_3}(y_2) = \frac{4}{3}$, for $X^{\geqslant}_3 = \{y_1, y_2\}$ and $X'^{\geqslant}_3 = \{y_1, y_2, y_3\}$. Now, let us calculate Bayes factors for object $y_2$ and unions of classes $X^{\geqslant}_2, X^{\geqslant}_3$. We have $B^P_{X^{\geqslant}_3}(y_2) = \frac{4}{3} > B^P_{X^{\geqslant}_2}(y_2) = \frac{1}{2}$. This shows that measure $B^P_{X^{\geqslant}}(y)$ does not have property (m3). Finally, let us notice that $B^P_{X^{\geqslant}_3}(y_2) = \frac{4}{3} > B^P_{X^{\geqslant}_3}(y_1) = \frac{2}{3}$. This proves, that measure $B^P_{X^{\geqslant}}(y)$ also does not have property (m4).

## 4. Monotonic Variable Consistency Indiscernibility-based Rough Set Approaches

Our motivation for proposing Variable Consistency Indiscernibility-based Rough Set Approaches (VC-IRSA) comes from the need of ensuring monotonicity of lower approximations w.r.t. set of attributes. Due to the definition of the upper approximation based on complementarity, w.r.t. the lower approximation, this monotonicity property also concerns the upper approximation. The main difference between VC-IRSA and VPRS [26,28,29], RB model [25] and PRS [14] is that in VC-IRSA one considers for inclusion to $P$-lower approximations only these objects which belong to the approximated set (definitions (2) and (4)). In VPRS, RB model and PRS whole granules are included to $P$-lower approximations (definitions (1) and (3)). Remark that a granule included in a $P$-lower approximation may be composed of some inconsistent objects. After enlarging




| object | $a_1$ | $a_2$ | decision class |
|---|---|---|---|
| $y_1$ | 3 | 3 | $X_3$ |
| $y_2$ | 2 | 2 | $X_3$ |
| $y_3$ | 1 | 3 | $X_3$ |
| $y_4$ | 3 | 1 | $X_2$ |
| $y_5$ | 4 | 4 | $X_1$ |

**Fig. 2.** Exemplary set of objects described by set $P$ of gain-type criteria $a_1$ and $a_2$.

set $P$ of attributes to $P' \supset P$, some of $P$-indiscernible and inconsistent objects may become $P'$-discernible and thus consistent, so, if we would like to preserve monotonicity of lower approximations, then we should keep in the $P'$-lower approximation the $P'$-discernible objects that do not belong to the approximated set. This, is not reasonable, however. Motivated by this remark, we consider only lower approximations defined according to (2) or (4).

Below, we introduce new consistency measures for VC-IRSA. We also present theorems concerning monotonicity properties of these measures. Proofs of all the theorems are given in the Appendix.

As it was already mentioned in Section 2, monotonicity properties of a consistency measure used in the definition of the $P$-lower approximation imply monotonicity properties of this approximation.

### 4.1. Consistency measure $\epsilon$

The first consistency measure that we consider in VC-IRSA is a cost-type measure $\epsilon^P_{X_i}(y)$. For $P \subseteq C, X_i, \neg X_i \subseteq U$, where $\neg X_i = U - X_i$, $y \in U$, it is defined as

$$\epsilon^P_{X_i}(y) = \frac{|I_P(y) \cap \neg X_i|}{|\neg X_i|}. \tag{17}$$

In the numerator of (17) there is the number of objects in $U$ that do not belong to class $X_i$ and are indiscernible with object $y$. In the denominator of (17) there is the number of objects in $U$ that do not belong to class $X_i$. The ratio $\epsilon^P_{X_i}(y)$ is an estimate of conditional probability $Pr(x \in I_P(y) | x \in \neg X_i)$, called also a catch-all likelihood [6]. This measure is an estimate of probability that object $y$ belongs to granule $I_P(y)$ given that it does not belong to class $X_i$. It may result in low values of consistency measure $\epsilon^P_{X_i}(y)$ for classes $X_i$ that have low cardinality.

**Theorem 4.1.** *Measure $\epsilon^P_{X_i}(y)$ has property (m1), i.e., for all $P \subseteq P' \subseteq C, X_i \subseteq U, y \in U$:*

$$\epsilon^P_{X_i}(y) \geqslant \epsilon^{P'}_{X_i}(y).$$

**Theorem 4.2.** *Measure $\epsilon^P_{X_i}(y)$ has property (m2). More precisely, for all $P \subseteq C$, $X_i \subseteq U$, $X'_i = X_i \cup X^\Delta_i$, $X^\Delta_i \cap U = \emptyset$, $y \in U$:*

$$\epsilon^P_{X_i}(y) = \epsilon^P_{X'_i}(y).$$

Monotonic $P$-lower approximation of class $X_i$ defined according to (4) takes the form:

$$\underline{P}^{\beta_{X_i}}(X_i) = \{y \in X_i : \epsilon^P_{X_i}(y) \leqslant \beta_{X_i}\}, \tag{18}$$

where cost-threshold $\beta_{X_i} \in [0, 1]$ reflects the highest degree of consistency acceptable to include object $y$ in the $P$-lower approximation of class $X_i$.

**Theorem 4.3.** *Lower approximation defined according to (18) satisfies condition (6):*

$$\underline{P}(X_i) \subseteq \underline{P}^{\beta_{X_i}}(X_i).$$

### 4.2. Consistency measure $\epsilon'$

Another consistency measure that we consider in VC-IRSA is a cost-type measure $\epsilon'^{P}_{X_i}(y)$. For $P \subseteq C, X_i, \neg X_i \subseteq U$, where $\neg X_i = U - X_i$, $y \in U$, it is defined as

$$\epsilon'^{P}_{X_i}(y) = \frac{|I_P(y) \cap \neg X_i|}{|X_i|}. \tag{19}$$

In the numerator of (19) there is the number of objects in $U$ that do not belong to class $X_i$ and are indiscernible with object $y$. In the denominator of (19) there is the number of objects in $U$ that belong to class $X_i$. This measure represents the ratio of objects $z \in U$ that are counterexamples to the implication $z \in I_P(y)$ implies $z \in X_i$ to the total number of objects in $X_i$. It lacks the likelihood interpretation that we give for $\epsilon^{P}_{X_i}(y)$. It should be noticed that $\epsilon'^{P}_{X_i}(y)$ may have low values for classes $X_i$ that have high cardinality.

**Theorem 4.4.** *Measure $\epsilon'^{P}_{X_i}(y)$ has property (m1), i.e., for all $P \subseteq P' \subseteq C, X_i \subseteq U, y \in U$:*

$$\epsilon'^{P}_{X_i}(y) \geqslant \epsilon'^{P'}_{X_i}(y).$$

**Theorem 4.5.** *Measure $\epsilon'^{P}_{X_i}(y)$ has property (m2). More precisely, for all $P \subseteq C$, $X_i \subseteq U$, $X'_i = X_i \cup X_i^{\Delta}$, $X_i^{\Delta} \cap U = \emptyset$, $y \in U$:*

$$\epsilon'^{P}_{X_i}(y) = \epsilon'^{P}_{X'_i}(y).$$

Monotonic $P$-lower approximation of class $X_i$ defined according to (4) takes the form:

$$\underline{P}^{\beta'_{X_i}}(X_i) = \{y \in X_i : \epsilon'^{P}_{X_i}(y) \leqslant \beta'_{X_i}\}, \tag{20}$$

where cost-threshold $\beta'_{X_i} \in \left[0, \frac{|\neg X_i|}{|X_i|}\right]$ reflects the highest degree of consistency acceptable to include object $y$ in the $P$-lower approximation of class $X_i$.

**Theorem 4.6.** *Lower approximation defined according to (20) satisfies condition (6):*

$$\underline{P}(X_i) \subseteq \underline{P}^{\beta'_{X_i}}(X_i).$$

### 4.3. Consistency measure $\overline{\mu}$

A gain-type consistency measure that can be considered in VC-IRSA is measure $\overline{\mu}^{P}_{X_i}(y)$. For $P \subseteq C, X_i \subseteq U, y \in U$, it is defined as

$$\overline{\mu}^{P}_{X_i}(y) = \max_{R \subseteq P} \frac{|I_R(y) \cap X_i|}{|I_R(y)|}. \tag{21}$$

Consistency measure $\overline{\mu}^{P}_{X_i}(y)$ is calculated as a maximum rough membership to class $X_i$ over all subsets $R$ of the set of attributes $P$.

**Theorem 4.7.** *Measure $\overline{\mu}^{P}_{X_i}(y)$ has property (m1), i.e., for all $P \subseteq P' \subseteq C, X_i \subseteq U, y \in U$:*

$$\overline{\mu}^{P}_{X_i}(y) \leqslant \overline{\mu}^{P'}_{X_i}(y).$$

**Theorem 4.8.** *Measure $\overline{\mu}^{P}_{X_i}(y)$ has property (m2), i.e., for all $P \subseteq C$, $X_i \subseteq U$, $X'_i = X_i \cup X_i^{\Delta}$, $X_i^{\Delta} \cap U = \emptyset$, $y \in U$:*

$$\overline{\mu}^{P}_{X_i}(y) \leqslant \overline{\mu}^{P}_{X'_i}(y).$$

Monotonic $P$-lower approximation of class $X_i$ defined according to (2) takes the form:

$$\underline{P}^{\bar{\alpha}_{X_i}}(X_i) = \{y \in X_i : \overline{\mu}^{P}_{X_i}(y) \geqslant \bar{\alpha}_{X_i}\}, \tag{22}$$

where gain-threshold $\bar{\alpha}_{X_i} \in [0, 1]$ reflects the lowest degree of consistency acceptable to include object $y$ in the $P$-lower approximation of class $X_i$.

**Theorem 4.9.** *Lower approximation defined according to (22) satisfies condition (5):*

$$\underline{P}(X_i) \subseteq \underline{P}^{\bar{\alpha}_{X_i}}(X_i).$$

In [2], we also considered a gain-type consistency measure $\underline{\mu}^{P}_{X_i}(y)$ which is defined analogously to $\overline{\mu}^{P}_{X_i}(y)$. For $P \subseteq C, X_i \subseteq U$, $y \in U$:

**Table 1**
Monotonicity of consistency measures defined for VC-IRSA.

| Consistency measure | (m1) | (m2) |
|---|---|---|
| $\epsilon^P_{X_i}(y)$ | Yes | Yes |
| $\epsilon'^P_{X_i}(y)$ | Yes | Yes |
| $\overline{\mu}^P_{X_i}(y)$ | Yes | Yes |

$$\underline{\mu}^P_{X_i}(y) = \min_{R \supseteq P} \frac{|I_R(y) \cap X_i|}{|I_R(y)|}.$$

It appears that this measure also has properties (m1) and (m2). However, it was used to define the $P$-lower approximation together with $\overline{\mu}^P_{X_i}(y)$. We refrain from using $\underline{\mu}^P_{X_i}(y)$ alone in the definition of the $P$-lower approximation.

### 4.4. Summary

In this section, we proposed definitions of three measures that ensure monotonicity of VC-IRSA. In Section 4.1 consistency measure $\epsilon$ was introduced. This measure has the meaning of a likelihood that an object is not a member of the considered class, given that it belongs to a granule of indiscernible objects. Such a kind of likelihood is sometimes called a catch-all likelihood. In Section 4.2 consistency measure $\epsilon'$ was introduced. This measure can be seen as complementary to measure $\epsilon$. They differ only by denominator. Monotonic measure defined in Section 4.3 involves rough membership measure $\mu$. It requires calculation of $\mu$ over all subsets of $P \subseteq C$. For all of these measures, we checked monotonicity properties (m1) and (m2). The results are summarized in Table 1.

## 5. Monotonic Variable Consistency Dominance-based Rough Set Approaches

We reformulate definitions of monotonic approaches presented in Section 4, replacing indiscernibility relation by dominance relation. Precisely, instead of granule $I_P(y)$, we use positive dominance cone $D^+_P(y)$ or negative dominance cone $D^-_P(y)$, and instead of decision class $X_i$, we consider upward union of decision classes $X_i^{\geqslant}$ or downward union of decision classes $X_i^{\leqslant}$. We also present theorems concerning monotonicity properties of the introduced consistency measures. Proofs of all the theorems are given in the Appendix.

As it was already mentioned in Section 2, monotonicity properties of a consistency measure used in the definition of the $P$-lower approximation imply monotonicity properties of this approximation.

### 5.1. Consistency measure $\epsilon$

Cost-type consistency measures $\epsilon^P_{X_i^{\geqslant}}(y)$ and $\epsilon^P_{X_i^{\leqslant}}(y)$, for $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant}, X_{i-1}^{\leqslant}, X_{i+1}^{\geqslant} \subseteq U$, $y \in U$, are defined as

$$\epsilon^P_{X_i^{\geqslant}}(y) = \frac{|D^+_P(y) \cap X_{i-1}^{\leqslant}|}{|X_{i-1}^{\leqslant}|}, \quad \epsilon^P_{X_i^{\leqslant}}(y) = \frac{|D^-_P(y) \cap X_{i+1}^{\geqslant}|}{|X_{i+1}^{\geqslant}|}. \tag{23}$$

Consistency measure $\epsilon^P_{X_i^{\geqslant}}(y)$ (or $\epsilon^P_{X_i^{\leqslant}}(y)$) can be interpreted as an estimate of conditional probability that object $y$ belongs to the considered dominance cone given that it does not belong to the considered union. In other words, it is the number of objects in the dominance cone of object $y$ that do not belong to the considered union of classes, divided by the number of all those objects that do not belong to the considered union of classes. Analogously to Section 4.1, measures $\epsilon^P_{X_i^{\geqslant}}(y)$ and $\epsilon^P_{X_i^{\leqslant}}(y)$ can be interpreted as catch-all likelihoods.

**Theorem 5.1.** *Measures $\epsilon^P_{X_i^{\geqslant}}(y)$ and $\epsilon^P_{X_i^{\leqslant}}(y)$ have property (m1), i.e., for all $P \subseteq P' \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:*

$$\epsilon^P_{X_i^{\geqslant}}(y) \geqslant \epsilon^{P'}_{X_i^{\geqslant}}(y), \quad \epsilon^P_{X_i^{\leqslant}}(y) \geqslant \epsilon^{P'}_{X_i^{\leqslant}}(y).$$

**Theorem 5.2.** *Measure $\epsilon^P_{X_i^{\geqslant}}(y)$ has property (m2). More precisely, for all $P \subseteq C$, $X_i^{\geqslant} \subseteq U$, $X_i'^{\geqslant} = X_i^{\geqslant} \cup X_i^{\Delta\geqslant}$, $X_i^{\Delta\geqslant} \cap U = \emptyset$, $y \in U$:*

$$\epsilon^P_{X_i^{\geqslant}}(y) = \epsilon^P_{X_i'^{\geqslant}}(y).$$

**Theorem 5.3.** *Measure $\epsilon^P_{X_i^{\leqslant}}(y)$ has property (m2). More precisely, for all $P \subseteq C$, $X_i^{\leqslant} \subseteq U$, $X_i'^{\leqslant} = X_i^{\leqslant} \cup X_i^{\Delta\leqslant}$, $X_i^{\Delta\leqslant} \cap U = \emptyset$, $y \in U$:*

$$\epsilon^P_{X_i^{\leqslant}}(y) = \epsilon^P_{X_i'^{\leqslant}}(y).$$

**Theorem 5.4.** *Measures $\epsilon^P_{X_i^{\geqslant}}(y)$ and $\epsilon^P_{X_i^{\leqslant}}(y)$ have property (m4), i.e., for all $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:*

$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow \epsilon^P_{X_i^{\geqslant}}(y_1) \leqslant \epsilon^P_{X_i^{\geqslant}}(y_2),$$
$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow \epsilon^P_{X_i^{\leqslant}}(y_1) \geqslant \epsilon^P_{X_i^{\leqslant}}(y_2).$$

Unfortunately, measures $\epsilon^P_{X_i^{\geqslant}}(y)$ and $\epsilon^P_{X_i^{\leqslant}}(y)$ do not have property (m3). More precisely, for all $P \subseteq C$, $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, $j \leqslant i$, $y \in U$, measure $\epsilon^P_{X_i^{\geqslant}}(y)$ is not monotonically non-increasing w.r.t. set of objects $X_i^{\geqslant}$, and for all $P \subseteq C$, $X_i^{\leqslant} \subseteq X_j^{\leqslant} \subseteq U$, $j \geqslant i$, $y \in U$, measure $\epsilon^P_{X_i^{\leqslant}}(y)$ is not monotonically non-increasing w.r.t. set of objects $X_i^{\leqslant}$. This can be illustrated by the following example. We have $P = \{a_1\}$, $X_1 = \{y_1\}$, $X_2 = \{y_2\}$, $X_3 = \{y_3\}$, where $f(y_1, a_1) = 3$, $f(y_2, a_1) = 1$, $f(y_3, a_1) = 2$. Moreover, let us assume that attribute $a_1$ is gain-type and decision classes are ordered such that class $X_3$ is better than $X_2$, which is better than $X_1$. We have, $\epsilon^P_{X_3^{\geqslant}}(y_3) = \frac{1}{2} < \epsilon^P_{X_2^{\geqslant}}(y_3) = 1$. The same can be shown for downward unions.

In order to ensure property (m3), in Sections 5.2 and 5.3 we introduce two possible modifications of measures $\epsilon^P_{X_i^{\geqslant}}(y)$ and $\epsilon^P_{X_i^{\leqslant}}(y)$.

### 5.2. Consistency measure $\epsilon^*$

Cost-type consistency measures $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ and $\epsilon^{*P}_{X_i^{\leqslant}}(y)$, for $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$, are defined as

$$\epsilon^{*P}_{X_i^{\geqslant}}(y) = \max_{j \leqslant i} \epsilon^P_{X_j^{\geqslant}}(y), \tag{24}$$
$$\epsilon^{*P}_{X_i^{\leqslant}}(y) = \max_{j \geqslant i} \epsilon^P_{X_j^{\leqslant}}(y). \tag{25}$$

Measure $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ (or $\epsilon^{*P}_{X_i^{\leqslant}}(y)$) is defined as a maximal value of measure $\epsilon^P_{X_i^{\geqslant}}(y)$ ($\epsilon^P_{X_i^{\leqslant}}(y)$) over all unions of decision classes which contain considered union $X_i^{\geqslant}$ ($X_i^{\leqslant}$). Measures $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ and $\epsilon^{*P}_{X_i^{\leqslant}}(y)$ satisfy all monotonicity properties of $\epsilon^P_{X_i^{\geqslant}}(y)$ and $\epsilon^P_{X_i^{\leqslant}}(y)$, respectively. Moreover, as we show below, they have also monotonicity property (m3).

**Theorem 5.5.** *Measures $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ and $\epsilon^{*P}_{X_i^{\leqslant}}(y)$ have property (m1), i.e., for all $P \subseteq P' \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:*

$$\epsilon^{*P}_{X_i^{\geqslant}}(y) \geqslant \epsilon^{*P'}_{X_i^{\geqslant}}(y), \quad \epsilon^{*P}_{X_i^{\leqslant}}(y) \geqslant \epsilon^{*P'}_{X_i^{\leqslant}}(y).$$

**Theorem 5.6.** *Measure $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ has property (m2). More precisely, for all $P \subseteq C$, $X_i^{\geqslant} \subseteq U$, $X_i'^{\geqslant} = X_i^{\geqslant} \cup X_i^{\Delta\geqslant}$, $X_i^{\Delta\geqslant} \cap U = \emptyset$, $y \in U$:*

$$\epsilon^{*P}_{X_i^{\geqslant}}(y) = \epsilon^{*P}_{X_i'^{\geqslant}}(y).$$

**Theorem 5.7.** *Measure $\epsilon^{*P}_{X_i^{\leqslant}}(y)$ has property (m2). More precisely, for all $P \subseteq C$, $X_i^{\leqslant} \subseteq U$, $X_i'^{\leqslant} = X_i^{\leqslant} \cup X_i^{\Delta\leqslant}$, $X_i^{\Delta\leqslant} \cap U = \emptyset$, $y \in U$:*

$$\epsilon^{*P}_{X_i^{\leqslant}}(y) = \epsilon^{*P}_{X_i'^{\leqslant}}(y).$$

**Theorem 5.8.** *Measure $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ has property (m3), i.e., for all $P \subseteq C$, $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, $j \leqslant i$, $y \in U$:*

$$\epsilon^{*P}_{X_i^{\geqslant}}(y) \geqslant \epsilon^{*P}_{X_j^{\geqslant}}(y).$$

**Theorem 5.9.** *Measure $\epsilon^{*P}_{X_i^{\leqslant}}(y)$ has property (m3), i.e., for all $P \subseteq C$, $X_i^{\leqslant} \subseteq X_j^{\leqslant} \subseteq U$, $j \geqslant i$, $y \in U$:*

$$\epsilon^{*P}_{X_i^{\leqslant}}(y) \geqslant \epsilon^{*P}_{X_j^{\leqslant}}(y).$$

**Theorem 5.10.** *Measures $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ and $\epsilon^{*P}_{X_i^{\leqslant}}(y)$ have property (m4), i.e., for all $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:*

$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow \epsilon^{*P}_{X_i^{\geqslant}}(y_1) \leqslant \epsilon^{*P}_{X_i^{\geqslant}}(y_2),$$
$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow \epsilon^{*P}_{X_i^{\leqslant}}(y_1) \geqslant \epsilon^{*P}_{X_i^{\leqslant}}(y_2).$$

Monotonic $P$-lower approximation of union of classes $X_i^{\geqslant}$, $X_i^{\leqslant}$ defined according to (4) takes the form:

$$\underline{P}^{\beta^*_{X_i^{\geqslant}}}(X_i^{\geqslant}) = \{y \in X_i^{\geqslant} : \epsilon^{*P}_{X_i^{\geqslant}}(y) \leqslant \beta^*_{X_i^{\geqslant}}\}, \tag{26}$$
$$\underline{P}^{\beta^*_{X_i^{\leqslant}}}(X_i^{\leqslant}) = \{y \in X_i^{\leqslant} : \epsilon^{*P}_{X_i^{\leqslant}}(y) \leqslant \beta^*_{X_i^{\leqslant}}\}, \tag{27}$$

where cost-threshold $\beta^*_{X_i^{\geqslant}}, \beta^*_{X_i^{\leqslant}} \in [0, 1]$ reflects the highest degree of consistency acceptable to include object $y$ in the $P$-lower approximation of union of classes $X_i^{\geqslant}$, $X_i^{\leqslant}$, respectively.

**Theorem 5.11.** *Lower approximations defined according to (26) and (27) satisfy condition (6):*

$$\underline{P}(X_i^{\geqslant}) \subseteq \underline{P}^{\beta^*_{X_i^{\geqslant}}}(X_i^{\geqslant}),$$
$$\underline{P}(X_i^{\leqslant}) \subseteq \underline{P}^{\beta^*_{X_i^{\leqslant}}}(X_i^{\leqslant}).$$

### 5.3. Consistency measure $\epsilon'$

Another way to overcome the lack of property (m3) of $\epsilon^P_{X_i^{\geqslant}}(y)$ and $\epsilon^P_{X_i^{\leqslant}}(y)$ is to consider cost-type consistency measures $\epsilon'^P_{X_i^{\geqslant}}(y)$ and $\epsilon'^P_{X_i^{\leqslant}}(y)$. For $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant}, X_{i-1}^{\leqslant}, X_{i+1}^{\geqslant} \subseteq U$, $y \in U$, they are defined as

$$\epsilon'^P_{X_i^{\geqslant}}(y) = \frac{|D_P^+(y) \cap X_{i-1}^{\leqslant}|}{|X_i^{\geqslant}|}, \quad \epsilon'^P_{X_i^{\leqslant}}(y) = \frac{|D_P^-(y) \cap X_{i+1}^{\geqslant}|}{|X_i^{\leqslant}|}. \tag{28}$$

Consistency measure $\epsilon'^P_{X_i^{\geqslant}}(y)$ (or $\epsilon'^P_{X_i^{\leqslant}}(y)$) is defined as a ratio of the number of objects that belong both to dominance cone $D_P^+(y)$ ($D_P^-(y)$) and union $X_{i-1}^{\leqslant}$ ($X_{i+1}^{\geqslant}$), to the number of objects belonging to union $X_i^{\geqslant}$ ($X_i^{\leqslant}$). In other words, this measure represents the ratio of objects $z \in U$ that are counterexamples to the implication $z \in D_P^+(y)$ ($z \in D_P^-(y)$) implies $z \in X_i^{\geqslant}$ ($z \in X_i^{\leqslant}$) to the total number of objects in $X_i^{\geqslant}$ ($X_i^{\leqslant}$).

**Theorem 5.12.** *Measures $\epsilon'^P_{X_i^{\geqslant}}(y)$ and $\epsilon'^P_{X_i^{\leqslant}}(y)$ have property (m1), i.e., for all $P \subseteq P' \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:*

$$\epsilon'^P_{X_i^{\geqslant}}(y) \geqslant \epsilon'^{P'}_{X_i^{\geqslant}}(y), \quad \epsilon'^P_{X_i^{\leqslant}}(y) \geqslant \epsilon'^{P'}_{X_i^{\leqslant}}(y).$$

**Theorem 5.13.** *Measure $\epsilon'^P_{X_i^{\geqslant}}(y)$ has property (m2). More precisely, for all $P \subseteq C$, $X_i^{\geqslant} \subseteq U$, $X_i'^{\geqslant} = X_i^{\geqslant} \cup X_i^{\Delta\geqslant}$, $X_i^{\Delta\geqslant} \cap U = \emptyset$, $y \in U$:*

$$\epsilon'^P_{X_i^{\geqslant}}(y) > \epsilon'^P_{X_i'^{\geqslant}}(y).$$

**Theorem 5.14.** *Measure $\epsilon'^P_{X_i^{\leqslant}}(y)$ has property (m2). More precisely, for all $P \subseteq C$, $X_i^{\leqslant} \subseteq U$, $X_i'^{\leqslant} = X_i^{\leqslant} \cup X_i^{\Delta\leqslant}$, $X_i^{\Delta\leqslant} \cap U = \emptyset$, $y \in U$:*

$$\epsilon'^P_{X_i^{\leqslant}}(y) > \epsilon'^P_{X_i'^{\leqslant}}(y).$$

**Theorem 5.15.** *Measure $\epsilon'^P_{X_i^{\geqslant}}(y)$ has property (m3), i.e., for all $P \subseteq C$, $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, $j \leqslant i$, $y \in U$:*

$$\epsilon'^P_{X_i^{\geqslant}}(y) \geqslant \epsilon'^P_{X_j^{\geqslant}}(y).$$

**Theorem 5.16.** *Measure $\epsilon'^P_{X_i^{\leqslant}}(y)$ has property (m3), i.e., for all $P \subseteq C$, $X_i^{\leqslant} \subseteq X_j^{\leqslant} \subseteq U$, $j \geqslant i$, $y \in U$:*

$$\epsilon'^P_{X_i^{\leqslant}}(y) \geqslant \epsilon'^P_{X_j^{\leqslant}}(y).$$

**Theorem 5.17.** *Measures $\epsilon'^P_{X_i^{\geqslant}}(y)$ and $\epsilon'^P_{X_i^{\leqslant}}(y)$ have property (m4), i.e., for all $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:*

$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow \epsilon'^P_{X_i^{\geqslant}}(y_1) \leqslant \epsilon'^P_{X_i^{\geqslant}}(y_2),$$
$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow \epsilon'^P_{X_i^{\leqslant}}(y_1) \geqslant \epsilon'^P_{X_i^{\leqslant}}(y_2).$$

Measure $\epsilon'^P_{X_i^{\geqslant}}(y)$ (or $\epsilon'^P_{X_i^{\leqslant}}(y)$) has different interpretation from consistency measures $\epsilon^P_{X_i^{\geqslant}}(y)$ ($\epsilon^P_{X_i^{\leqslant}}(y)$) and $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ ($\epsilon^{*P}_{X_i^{\leqslant}}(y)$). It lacks likelihood explanation that is appropriate for the other two measures. It relates two rather antagonistic concepts. According to the definition of $\epsilon'^P_{X_i^{\geqslant}}(y)$ ($\epsilon'^P_{X_i^{\leqslant}}(y)$), the number of objects in the dominance cone of considered object $y$ that do not belong to the considered union of classes is divided by the cardinality of the considered union of classes. This may result in low values of consistency measure $\epsilon'^P_{X_i^{\geqslant}}(y)$ ($\epsilon'^P_{X_i^{\leqslant}}(y)$) for unions of classes $X_i^{\geqslant}$ ($X_i^{\leqslant}$) that have high cardinality.

Monotonic $P$-lower approximation of union of classes $X_i^{\geqslant}$, $X_i^{\leqslant}$ defined according to (4) takes the form:

$$\underline{P}^{\beta'_{X_i^{\geqslant}}}(X_i^{\geqslant}) = \{y \in X_i^{\geqslant} : \epsilon'^P_{X_i^{\geqslant}}(y) \leqslant \beta'_{X_i^{\geqslant}}\}, \tag{29}$$

$$\underline{P}^{\beta'_{X_i^{\leqslant}}}(X_i^{\leqslant}) = \{y \in X_i^{\leqslant} : \epsilon'^P_{X_i^{\leqslant}}(y) \leqslant \beta'_{X_i^{\leqslant}}\}, \tag{30}$$

where cost-threshold $\beta'_{X_i^{\geqslant}} \in \left[0, \frac{|X_{i-1}^{\leqslant}|}{|X_i^{\geqslant}|}\right]$, $\beta'_{X_i^{\leqslant}} \in \left[0, \frac{|X_{i+1}^{\geqslant}|}{|X_i^{\leqslant}|}\right]$ reflects the highest degree of consistency acceptable to include object $y$ in the $P$-lower approximation of union of classes $X_i^{\geqslant}$, $X_i^{\leqslant}$, respectively.

**Theorem 5.18.** *Lower approximations defined according to (29) and (30) satisfy condition (6):*

$$\underline{P}(X_i^{\geqslant}) \subseteq \underline{P}^{\beta'_{X_i^{\geqslant}}}(X_i^{\geqslant}),$$
$$\underline{P}(X_i^{\leqslant}) \subseteq \underline{P}^{\beta'_{X_i^{\leqslant}}}(X_i^{\leqslant}).$$

*5.4. Consistency measure $\overline{\mu}$*

For $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$, we also consider the following gain-type consistency measures:

$$\overline{\mu}^P_{X_i^{\geqslant}}(y) = \max_{\substack{R \subseteq P,\\ z \in D_R^-(y) \cap X_i^{\geqslant}}} \frac{|D_R^+(z) \cap X_i^{\geqslant}|}{|D_R^+(z)|}, \tag{31}$$

$$\overline{\mu}^P_{X_i^{\leqslant}}(y) = \max_{\substack{R \subseteq P,\\ z \in D_R^+(y) \cap X_i^{\leqslant}}} \frac{|D_R^-(z) \cap X_i^{\leqslant}|}{|D_R^-(z)|}. \tag{32}$$

Measure $\overline{\mu}^P_{X_i^{\geqslant}}(y)$ (or $\overline{\mu}^P_{X_i^{\leqslant}}(y)$) is defined as a maximum rough membership to union $X_i^{\geqslant}$ ($X_i^{\leqslant}$) over all subsets $R$ of the set of attributes $P$ and over all objects $z$ dominated by $y$ (dominating $y$) and belonging to $X_i^{\geqslant}$ ($X_i^{\leqslant}$). Comparing the above definitions with the analogous definition (21) presented for VC-IRSA, one can easily observe that they have a new ingredient – the maximum is calculated not only over all subsets $R$ of $P$ but also over all objects belonging to the intersection of the particular dominance cone of object $y$ and the considered union of decision classes. Such a formulation ensures monotonicity property (m4), which is proved later in this section.

**Theorem 5.19.** *Measures $\overline{\mu}^P_{X_i^{\geqslant}}(y)$ and $\overline{\mu}^P_{X_i^{\leqslant}}(y)$ have property (m1), i.e., for all $P \subseteq P' \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:*

$$\overline{\mu}^P_{X_i^{\geqslant}}(y) \leqslant \overline{\mu}^{P'}_{X_i^{\geqslant}}(y),$$
$$\overline{\mu}^P_{X_i^{\leqslant}}(y) \leqslant \overline{\mu}^{P'}_{X_i^{\leqslant}}(y).$$

**Theorem 5.20.** *Measure $\overline{\mu}^P_{X_i^{\geqslant}}(y)$ has property (m2), i.e., for all $P \subseteq C$, $X_i^{\geqslant} \subseteq U$, $X_i'^{\geqslant} = X_i^{\geqslant} \cup X_i^{\Delta\geqslant}$, $X_i^{\Delta\geqslant} \cap U = \emptyset$, $y \in U$:*

$$\overline{\mu}^P_{X_i^{\geqslant}}(y) \leqslant \overline{\mu}^P_{X_i'^{\geqslant}}(y).$$

**Theorem 5.21.** *Measure $\overline{\mu}^P_{X_i^{\leqslant}}(y)$ has property (m2), i.e., for all $P \subseteq C$, $X_i^{\leqslant} \subseteq U$, $X_i'^{\leqslant} = X_i^{\leqslant} \cup X_i^{\Delta\leqslant}$, $X_i^{\Delta\leqslant} \cap U = \emptyset$, $y \in U$:*

$$\overline{\mu}^P_{X_i^{\leqslant}}(y) \leqslant \overline{\mu}^P_{X_i'^{\leqslant}}(y).$$

**Theorem 5.22.** *Measure $\overline{\mu}^P_{X_i^{\geqslant}}(y)$ has property (m3), i.e., for all $P \subseteq C$, $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, $j \leqslant i$, $y \in U$:*

$$\overline{\mu}^P_{X_i^{\geqslant}}(y) \leqslant \overline{\mu}^P_{X_j^{\geqslant}}(y).$$

**Theorem 5.23.** *Measure $\overline{\mu}^P_{X_i^{\leqslant}}(y)$ has property (m3), i.e, for all $P \subseteq C$, $X_i^{\leqslant} \subseteq X_j^{\leqslant} \subseteq U$, $j \geqslant i$, $y \in U$:*

$$\overline{\mu}^P_{X_i^{\leqslant}}(y) \leqslant \overline{\mu}^P_{X_j^{\leqslant}}(y).$$

**Theorem 5.24.** *Measures $\overline{\mu}^P_{X_i^{\geqslant}}(y)$ and $\overline{\mu}^P_{X_i^{\leqslant}}(y)$ have property (m4), i.e., for all $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:*

$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow \overline{\mu}^P_{X_i^{\geqslant}}(y_1) \geqslant \overline{\mu}^P_{X_i^{\geqslant}}(y_2),$$
$$\forall y_1, y_2 \in U : y_1 D_P y_2 \Rightarrow \overline{\mu}^P_{X_i^{\leqslant}}(y_1) \leqslant \overline{\mu}^P_{X_i^{\leqslant}}(y_2).$$

Monotonic $P$-lower approximation of union of classes $X_i^{\geqslant}$, $X_i^{\leqslant}$ defined according to (2) takes the form:

$$\underline{P}^{\bar{\alpha}_{X_i^{\geqslant}}}(X_i^{\geqslant}) = \{y \in X_i^{\geqslant} : \overline{\mu}^P_{X_i^{\geqslant}}(y) \geqslant \bar{\alpha}_{X_i^{\geqslant}}\}, \tag{33}$$

$$\underline{P}^{\bar{\alpha}_{X_i^{\leqslant}}}(X_i^{\leqslant}) = \{y \in X_i^{\leqslant} : \overline{\mu}^P_{X_i^{\leqslant}}(y) \geqslant \bar{\alpha}_{X_i^{\leqslant}}\}, \tag{34}$$

where gain-threshold $\bar{\alpha}_{X_i^{\geqslant}}, \bar{\alpha}_{X_i^{\leqslant}} \in [0, 1]$ reflects the lowest degree of consistency acceptable to include object $y$ in the $P$-lower approximation of union of classes $X_i^{\geqslant}$, $X_i^{\leqslant}$, respectively.

**Table 2**
Monotonicity of consistency measures defined for VC-DRSA.

| Consistency measure | (m1) | (m2) | (m3) | (m4) |
|---|---|---|---|---|
| $\epsilon^{P}_{X_i^{\geqslant}}(y)$, $\epsilon^{P}_{X_i^{\leqslant}}(y)$ | Yes | Yes | No | Yes |
| $\epsilon^{*P}_{X_i^{\geqslant}}(y)$, $\epsilon^{*P}_{X_i^{\leqslant}}(y)$ | Yes | Yes | Yes | Yes |
| $\epsilon'^{P}_{X_i^{\geqslant}}(y)$, $\epsilon'^{P}_{X_i^{\leqslant}}(y)$ | Yes | Yes | Yes | Yes |
| $\overline{\mu}^{P}_{X_i^{\geqslant}}(y)$, $\overline{\mu}^{P}_{X_i^{\leqslant}}(y)$ | Yes | Yes | Yes | Yes |

**Theorem 5.25.** *Lower approximations defined according to (33) and (34) satisfy condition (5):*

$$\underline{P}(X_i^{\geqslant}) \subseteq \underline{P}^{\tilde{\alpha}_{X_i^{\geqslant}}}(X_i^{\geqslant}),$$
$$\underline{P}(X_i^{\leqslant}) \subseteq \underline{P}^{\tilde{\alpha}_{X_i^{\leqslant}}}(X_i^{\leqslant}).$$

In [2], we considered gain-type consistency measures $\underline{\mu}^{P}_{X_i^{\geqslant}}(y)$ and $\underline{\mu}^{P}_{X_i^{\leqslant}}(y)$, which are defined analogously to $\overline{\mu}^{P}_{X_i^{\geqslant}}(y)$ and $\overline{\mu}^{P}_{X_i^{\leqslant}}(y)$. For $P \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$:

$$\underline{\mu}^{P}_{X_i^{\geqslant}}(y) = \min_{\substack{R \supseteq P,\\ z \in D_R^{+}(y) \cap X_i^{\geqslant}}} \frac{|D_R^{+}(z) \cap X_i^{\geqslant}|}{|D_R^{+}(z)|}, \quad \underline{\mu}^{P}_{X_i^{\leqslant}}(y) = \min_{\substack{R \supseteq P,\\ z \in D_R^{-}(y) \cap X_i^{\leqslant}}} \frac{|D_R^{-}(z) \cap X_i^{\leqslant}|}{|D_R^{-}(z)|}.$$

It appears that these measures have properties (m1) and (m4) while they do not have properties (m2) and (m3). Therefore, we refrained from using them in the definition of the $P$-lower approximation.

### 5.5. Summary

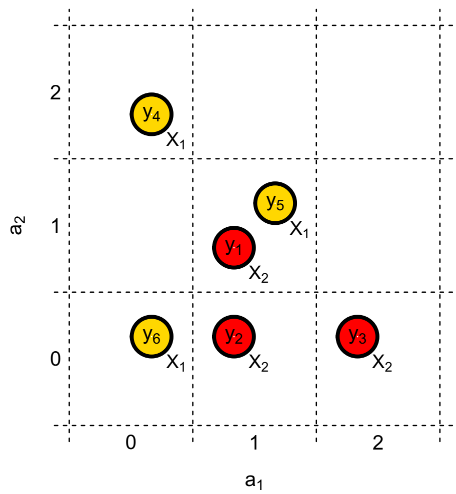
In this section, we introduced several consistency measures for VC-DRSA. Their properties are summarized in Table 2. Remark that $\epsilon^{P}_{X_i^{\geqslant}}(y)$ and $\epsilon^{P}_{X_i^{\leqslant}}(y)$ are the only measures missing desirable property (m3). Therefore, two possible modifications of these measures, denoted by $\epsilon^{*P}_{X_i^{\geqslant}}(y)$, $\epsilon^{*P}_{X_i^{\leqslant}}(y)$ and $\epsilon'^{P}_{X_i^{\geqslant}}(y)$, $\epsilon'^{P}_{X_i^{\leqslant}}(y)$, were further investigated.

## 6. Illustrative example

Let us consider VC-IRSA and the set of objects shown in Fig. 3.

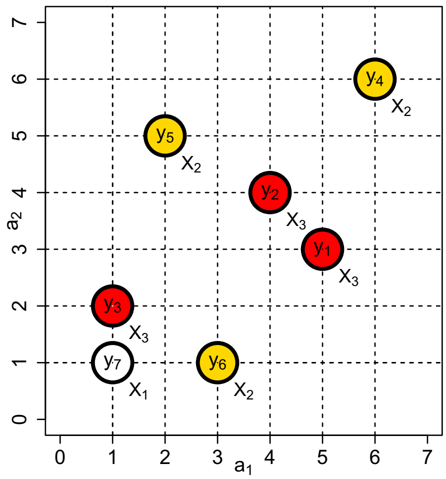
First, let us determine the $P$-lower approximation of class $X_2$ using definition (18), for $\beta_{X_2} = 0$. We can observe that $\underline{P}^0(X_2) = \{y_2, y_3\}$. Object $y_1$ is not included in $\underline{P}^0(X_2)$ because $\epsilon^{P}_{X_2}(y_1) = \frac{1}{3}$. We can also notice that $\epsilon^{\{a_1\}}_{X_2}(y_2) = \epsilon^{\{a_2\}}_{X_2}(y_2) = \frac{1}{3}$, while $\epsilon^{P}_{X_2}(y_2) = 0$. This illustrates property (m1) of measure $\epsilon^{P}_{X_i}(y)$. In order to exemplify property (m2) of measure $\epsilon^{P}_{X_i}(y)$, we extend class $X_2$ by adding to this class new object $y_7$, with the following description: $a_1 = 1$, $a_2 = 1$. One can easily verify that values of measure $\epsilon^{P}_{X_i}(y)$ for class $X_2$ and objects $y_1$, $y_2$ and $y_3$ do not change.

Second, let us calculate the $P$-lower approximation of class $X_2$ using definition (20), for $\beta'_{X_2} = 0$. We can notice that $\underline{P}^0(X_2) = \{y_2, y_3\}$. Object $y_1$ is not included in $\underline{P}^0(X_2)$ because $\epsilon'^{P}_{X_2}(y_1) = \frac{1}{3}$. We can also notice that $\epsilon'^{\{a_1\}}_{X_2}(y_2) = \epsilon'^{\{a_2\}}_{X_2}(y_2) = \frac{1}{3}$,




| object | $a_1$ | $a_2$ | decision class |
|---|---|---|---|
| $y_1$ | 1 | 1 | $X_2$ |
| $y_2$ | 1 | 0 | $X_2$ |
| $y_3$ | 2 | 0 | $X_2$ |
| $y_4$ | 0 | 2 | $X_1$ |
| $y_5$ | 1 | 1 | $X_1$ |
| $y_6$ | 0 | 0 | $X_1$ |

**Fig. 3.** Exemplary set of objects described by set $P$ of attributes $a_1$ and $a_2$.




| object | $a_1$ | $a_2$ | decision class |
|---|---|---|---|
| $y_1$ | 5 | 3 | $X_3$ |
| $y_2$ | 4 | 4 | $X_3$ |
| $y_3$ | 1 | 2 | $X_3$ |
| $y_4$ | 6 | 6 | $X_2$ |
| $y_5$ | 2 | 5 | $X_2$ |
| $y_6$ | 3 | 1 | $X_2$ |
| $y_7$ | 1 | 1 | $X_1$ |

**Fig. 4.** Exemplary set of objects described by set $P$ of gain-type criteria $a_1$ and $a_2$.

while $\epsilon'^{P}_{X_2}(y_2) = 0$. This illustrates property (m1) of measure $\epsilon'^{P}_{X_i}(y)$. In order to exemplify property (m2) of measure $\epsilon'^{P}_{X_i}(y)$, we extend class $X_2$ by adding to this class new object $y_7$, with the following description: $a_1 = 1, a_2 = 1$. One can easily verify that values of measure $\epsilon'^{P}_{X_i}(y)$ for class $X_2$ and objects $y_1$, $y_2$ and $y_3$ decrease.

Third, let us consider measure $\overline{\mu}^{P}_{X_i}(y)$ and definition (22). For $\bar{\alpha}_{X_2} = \frac{2}{3}$ we have $\underline{P}^{\frac{2}{3}}(X_2) = \{y_1, y_2, y_3\}$. It is worth noting that because $\mu^{\{a_1\}}_{X_2}(y_1) = \frac{2}{3}$, $\mu^{\{a_2\}}_{X_2}(y_1) = \frac{1}{2}$ and $\mu^{P}_{X_2}(y_1) = \frac{1}{2}$, then $\overline{\mu}^{P}_{X_2}(y_1) = \frac{2}{3}$. Thus, we can observe that measure $\overline{\mu}^{P}_{X_i}(y)$ has property (m1). Let us now extend the set of objects from Fig. 3 with object $y_7 \in X_2$, for which $a_1 = 1, a_2 = 1$, as we did in the first example. Thus, we obtain $\overline{\mu}^{P}_{X_2}(y_1) = \frac{3}{4} \geqslant \frac{2}{3}$. This shows that measure $\overline{\mu}^{P}_{X_i}(y)$ has property (m2).

Now, let us apply VC-DRSA to the set of objects presented in Fig. 4. Class $X_3$ is better than class $X_2$, which is better than class $X_1$. Thus, we may distinguish two upward unions of decision classes: $X_3^{\geqslant}$, $X_2^{\geqslant}$, and two downward unions of decision classes: $X_1^{\leqslant}$, $X_2^{\leqslant}$.

First, let us determine the $P$-lower approximation of union $X_3^{\geqslant}$ using definition (26), for $\beta^{*}_{X_3^{\geqslant}} = \frac{1}{3}$. We can observe that $\underline{P}^{\frac{1}{3}}(X_3^{\geqslant}) = \{y_1, y_2\}$. Object $y_3$ is not included in $\underline{P}^{\frac{1}{3}}(X_3^{\geqslant})$ because $\epsilon^{*P}_{X_3^{\geqslant}}(y_3) = \max\{\frac{2}{4}, \frac{0}{1}\} = \frac{1}{2} > \frac{1}{3}$. We can notice that $\epsilon^{*\{a_1\}}_{X_3^{\geqslant}}(y_2) = \frac{1}{4}$, $\epsilon^{*\{a_2\}}_{X_3^{\geqslant}}(y_2) = \frac{1}{2}$, while $\epsilon^{*P}_{X_3^{\geqslant}}(y_2) = \frac{1}{4}$. This illustrates property (m1) of measure $\epsilon^{*P}_{X_i^{\geqslant}}(y)$. In order to exemplify property (m2) of measure $\epsilon^{*P}_{X_i^{\geqslant}}(y)$, we consider possible extension of union $X_3^{\geqslant}$ by new object $y_8 \in X_3$, with the following description: $a_1 = 5, a_2 = 6$. One can easily verify that in such a case values of measure $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ for extended union $X_3^{\geqslant}$ and objects $y_1, y_2, y_3$ do not change. Let us also observe that $\epsilon^{*P}_{X_3^{\geqslant}}(y_3) = \frac{1}{2}$, while $\epsilon^{*P}_{X_2^{\geqslant}}(y_3) = 0$. This illustrates property (m3) of measure $\epsilon^{*P}_{X_i^{\geqslant}}(y)$. Moreover, $\epsilon^{*P}_{X_3^{\geqslant}}(y_2) = \frac{1}{4}$, $\epsilon^{*P}_{X_3^{\geqslant}}(y_3) = \frac{1}{2}$ and $y_2 D_P y_3$ for $P = \{a_1, a_2\}$, which shows that measure $\epsilon^{*P}_{X_i^{\geqslant}}(y)$ has property (m4).

Second, let us calculate the $P$-lower approximation of union $X_3^{\geqslant}$ using definition (29), for $\beta'_{X_3^{\geqslant}} = \frac{1}{3}$. We have $\underline{P}^{\frac{1}{3}}(X_3^{\geqslant}) = \{y_1, y_2\}$. Object $y_3$ is not included in $\underline{P}^{\frac{1}{3}}(X_3^{\geqslant})$ because $\epsilon'^{P}_{X_3^{\geqslant}}(y_3) = \frac{2}{3} > \frac{1}{3}$. We can observe that $\epsilon'^{\{a_1\}}_{X_3^{\geqslant}}(y_2) = \frac{1}{3}$, $\epsilon'^{\{a_2\}}_{X_3^{\geqslant}}(y_2) = \frac{2}{3}$, while $\epsilon'^{P}_{X_3^{\geqslant}}(y_2) = \frac{1}{3}$. This illustrates property (m1) of measure $\epsilon'^{P}_{X_i^{\geqslant}}(y)$. In order to exemplify property (m2) of measure $\epsilon'^{P}_{X_i^{\geqslant}}(y)$, we consider possible extension of union $X_3^{\geqslant}$ by new object $y_8 \in X_3$, having the following description: $a_1 = 5$, $a_2 = 6$. One can easily verify that in such a case values of measure $\epsilon'^{P}_{X_i^{\geqslant}}(y)$ for extended union $X_3^{\geqslant}$ and objects $y_1$, $y_2$, $y_3$ decrease, since $|X_3^{\geqslant}|$ increases. Let us also notice that $\epsilon'^{P}_{X_3^{\geqslant}}(y_3) = \frac{2}{3}$, while $\epsilon'^{P}_{X_2^{\geqslant}}(y_3) = 0$. This illustrates property (m3) of measure $\epsilon'^{P}_{X_i^{\geqslant}}(y)$. Moreover, $\epsilon'^{P}_{X_3^{\geqslant}}(y_2) = \frac{1}{3}$, $\epsilon'^{P}_{X_3^{\geqslant}}(y_3) = \frac{2}{3}$ and $y_2 D_P y_3$ for $P = \{a_1, a_2\}$, which shows that measure $\epsilon'^{P}_{X_i^{\geqslant}}(y)$ has property (m4).

Third, let us consider measure $\overline{\mu}^{P}_{X_i^{\geqslant}}(y)$ and corresponding definition (33). For $\bar{\alpha}_{X_3^{\geqslant}} = \frac{2}{3}$ we have $\underline{P}^{\frac{2}{3}}(X_3^{\geqslant}) = \{y_1, y_2\}$. It is worth noting that because $\overline{\mu}^{\{a_1\}}_{X_3^{\geqslant}}(y_1) = \mu^{\{a_1\}}_{X_3^{\geqslant}}(y_2) = \frac{2}{3}$, $\overline{\mu}^{\{a_2\}}_{X_3^{\geqslant}}(y_1) = \mu^{\{a_2\}}_{X_3^{\geqslant}}(y_3) = \frac{3}{5}$ and $\mu^{P}_{X_3^{\geqslant}}(y_3) = \frac{3}{5}$, then $\overline{\mu}^{P}_{X_3^{\geqslant}}(y_1) = \frac{2}{3}$. Thus, we can observe that measure $\overline{\mu}^{P}_{X_i^{\geqslant}}(y)$ has property (m1). In order to illustrate property (m2) of measure $\overline{\mu}^{P}_{X_i^{\geqslant}}(y)$, we consider possible extension of union $X_3^{\geqslant}$ by new object $y_8 \in X_3$, with the following description: $a_1 = 5$, $a_2 = 6$. Then, we have $\overline{\mu}^{P}_{X_3^{\geqslant}}(y_1) = \frac{3}{4} > \frac{2}{3}$. Let us also observe that $\overline{\mu}^{P}_{X_3^{\geqslant}}(y_3) = \frac{3}{5}$, while $\overline{\mu}^{P}_{X_2^{\geqslant}}(y_3) = 1$. This exemplifies property (m3) of measure $\overline{\mu}^{P}_{X_i^{\geqslant}}(y)$. Moreover, $\overline{\mu}^{P}_{X_3^{\geqslant}}(y_2) = \frac{2}{3} > \overline{\mu}^{P}_{X_3^{\geqslant}}(y_3) = \frac{3}{5}$, while $y_2 D_P y_3$ for $P = \{a_1, a_2\}$, which shows that measure $\overline{\mu}^{P}_{X_i^{\geqslant}}(y)$ has property (m4).

## 7. Final remarks and conclusions

In this paper, we have presented several definitions of monotonic Variable Consistency Rough Set Approaches that employ indiscernibility or dominance relation. We have stressed the importance of some monotonicity properties of the consistency measure used in the definition of a lower approximation. We have considered the following monotonicity properties: (m1) – monotonicity w.r.t. set of attributes, (m2) and (m3) – monotonicity w.r.t. set of objects (where (m2) corresponds to growing universe $U$ and (m3) to fixed universe $U$ with growing unions of decision classes), and (m4) – monotonicity w.r.t. dominance relation (for approaches based on dominance relation only).

We have proposed two types of measures enjoying the above monotonicity properties. The first type stems from consistency measure $\epsilon$, which is a catch-all likelihood measure. This consistency measure has a comprehensible probabilistic explanation. It has also a close relation with the Bayes factor and confirmation measure $l$. We proposed a kind of complementary measure to $\epsilon$ denoted by $\epsilon'$. One can observe that for $\epsilon$, there is a tendency of including relatively more objects to lower approximations when the approximated class or union of classes has low cardinality. On the other hand, one can observe that for $\epsilon'$, there is a tendency of including relatively more objects to lower approximations when the approximated class or union of classes has high cardinality. Both of these measures are directly applicable in VC-IRSA. Unfortunately, in the context of VC-DRSA, measure $\epsilon$ does not have property (m3). In order to overcome this problem, we introduced measure $\epsilon^*$ that involves a specific scheme of computation of consistency measure $\epsilon$ over supersets of the considered union of classes.

Monotonic measures of the second type stem from consistency measure $\mu$. They require to take into account all subsets of the set of considered attributes. Computation of lower approximations defined by means of monotonic measure $\overline{\mu}$ is an NP-hard problem, equivalent to induction of a set of all rules. On the other hand, computation of such approximations and rule induction can be combined, and thus the total time would be of the same order as the time for induction of all rules.

As a conclusion, we can recommend using consistency measure $\epsilon$ or $\epsilon'$ for VC-IRSA and consistency measure $\epsilon^*$ or $\epsilon'$ for VC-DRSA. These measures have all required monotonicity properties and are much less computationally intensive than the monotonic measures of the second type.

**Acknowledgement**

The first, third and fourth author wish to acknowledge financial support from the Ministry of Science and Higher Education, grant N N519 314435.

**Appendix.**

**Proof of Theorem 4.1.** From the definition of rough granules $I_P(y)$ and $I_{P'}(y)$, $P \subseteq P' \subseteq C$, $y \in U$,

$$I_P(y) \supseteq I_{P'}(y)$$

for $X_i, \neg X_i \subseteq U$ being both independent of sets of considered attributes $P$ and $P'$. This implies:

$$\frac{|I_P(y) \cap \neg X_i|}{|\neg X_i|} \geqslant \frac{|I_{P'}(y) \cap \neg X_i|}{|\neg X_i|} \iff \epsilon^P_{X_i}(y) \geqslant \epsilon^{P'}_{X_i}(y). \qquad \square$$

**Proof of Theorem 4.2.** Since new objects are introduced to class $X_i \subseteq U$, thus for all sets of objects $X_i \subseteq U$, $X'_i = X_i \cup X^\Delta_i$, where $X^\Delta_i \cap U = \emptyset$,

$$\neg X_i = \neg X'_i.$$

For all $P \subseteq C$, $y \in U$, this implies:

$$\frac{|I_P(y) \cap \neg X_i|}{|\neg X_i|} = \frac{|I'_P(y) \cap \neg X'_i|}{|\neg X'_i|} \iff \epsilon^P_{X_i}(y) = \epsilon^P_{X'_i}(y),$$

where $I'_P(y)$ denotes a set of objects indiscernible with object $y$ when considering set of attributes $P$ and universe $U \cup X^\Delta_i$. $\square$

**Proof of Theorem 4.3.** For each object $y \in X_i$, $I_P(y) \subseteq X_i$ iff $\epsilon^P_{X_i}(y) = 0$. $\square$

**Proof of Theorem 4.4.** Analogous to Proof of Theorem 4.1 for measure $\epsilon^P_{X_i}(y)$ – only the common denominators in fractions are changed from $|\neg X_i|$ to $|X_i|$. $\square$

**Proof of Theorem 4.5.** New objects are introduced to class $X_i \subseteq U$. Thus, for all sets of objects $X_i \subseteq U$, $X'_i = X_i \cup X^\Delta_i$, where $X^\Delta_i \cap U = \emptyset$,

$$\neg X_i = \neg X'_i, \quad |X_i| < |X'_i|.$$

This implies that for all $P \subseteq C$, $y \in U$:

$$\frac{|I_P(y) \cap \neg X_i|}{|X_i|} > \frac{|I'_P(y) \cap \neg X'_i|}{|X'_i|} \iff \epsilon'^P_{X_i}(y) > \epsilon'^P_{X'_i}(y),$$

where $I'_P(y)$ denotes a set of objects indiscernible with object $y$ when considering set of attributes $P$ and universe $U \cup X^\Delta_i$. $\square$

**Proof of Theorem 4.6.** For each object $y \in X_i$, $I_P(y) \subseteq X_i$ iff $\epsilon'^{P}_{X_i}(y) = 0$. $\square$

**Proof of Theorem 4.7.** For all $P \subseteq P' \subseteq C,\ X_i \subseteq U,\ y \in U$,

$$\overline{\mu}^{P}_{X_i}(y) = \max_{R \subseteq P} \frac{|I_R(y) \cap X_i|}{|I_R(y)|} \leqslant \max_{R \subseteq P'} \frac{|I_R(y) \cap X_i|}{|I_R(y)|} = \overline{\mu}^{P'}_{X_i}(y). \qquad \square$$

**Proof of Theorem 4.8.** Let us consider $P \subseteq C$, $X_i \subseteq U$, $X'_i = X_i \cup X^{\Delta}_i$, $X^{\Delta}_i \cap U = \emptyset$, $y \in U$. Since all new objects are added to class $X_i$, both numerator and denominator of fraction

$$\frac{|I_P(y) \cap X_i|}{|I_P(y)|} = \mu^{P}_{X_i}(y)$$

can increase only with the same number $k \geqslant 0$, equal to difference $|I'_P(y)| - |I_P(y)|$:

$$\frac{|I_P(y) \cap X_i| + k}{|I_P(y)| + k} = \frac{|I'_P(y) \cap X'_i|}{|I'_P(y)|} = \mu^{P}_{X'_i}(y),$$

where $I'_P(y)$ denotes a set of objects indiscernible with object $y$ when considering set of attributes $P$ and universe $U \cup X^{\Delta}_i$. Further, let us introduce the following notation: $a = |I_P(y) \cap X_i|$, $b = |I_P(y)|$, and let us notice that $a \leqslant b$. We can observe that

$$\mu^{P}_{X_i}(y) \leqslant \mu^{P}_{X'_i}(y), \tag{35}$$

which is proved in the following way:

$$\frac{a}{b} \leqslant \frac{a+k}{b+k} \iff a(b+k) \leqslant b(a+k) \iff ab + ak \leqslant ab + bk \iff ak \leqslant bk \iff a \leqslant b.$$

Thus,

$$\overline{\mu}^{P}_{X_i}(y) = \max_{R \subseteq P} \mu^{R}_{X_i}(y) \leqslant \max_{R \subseteq P} \mu^{R}_{X'_i}(y) = \overline{\mu}^{P}_{X'_i}(y). \qquad \square$$

**Proof of Theorem 4.9.** For each object $y \in X_i$, $I_P(y) \subseteq X_i$ iff $\overline{\mu}^{P}_{X_i}(y) = 1$. $\square$

**Proof of Theorem 5.1.** From the definition of dominance cones $D^{+}_{P}(y)$ and $D^{+}_{P'}(y)$, $P \subseteq P' \subseteq C$, $y \in U$,

$$D^{+}_{P}(y) \supseteq D^{+}_{P'}(y)$$

for $X^{\geqslant}_i, X^{\leqslant}_{i-1} \subseteq U$ being both independent of sets of considered attributes $P$ and $P'$. This implies:

$$\frac{|D^{+}_{P}(y) \cap X^{\leqslant}_{i-1}|}{|X^{\leqslant}_{i-1}|} \geqslant \frac{|D^{+}_{P'}(y) \cap X^{\leqslant}_{i-1}|}{|X^{\leqslant}_{i-1}|} \iff \epsilon^{P}_{X^{\geqslant}_i}(y) \geqslant \epsilon^{P'}_{X^{\geqslant}_i}(y).$$

The proof for downward union $X^{\leqslant}_i$ is analogical, but starts from the observation that for negative dominance cones $D^{-}_{P}(y)$ and $D^{-}_{P'}(y)$, $P \subseteq P' \subseteq C$, $y \in U$,

$$D^{-}_{P}(y) \supseteq D^{-}_{P'}(y). \qquad \square$$

**Proof of Theorem 5.2.** New objects are introduced to union of classes $X^{\geqslant}_i \subseteq U$. Thus, for all sets of objects $X^{\geqslant}_i \subseteq U$, $X'^{\geqslant}_i = X^{\geqslant}_i \cup X^{\Delta\geqslant}_i$, where $X^{\Delta\geqslant}_i \cap U = \emptyset$,

$$X^{\leqslant}_{i-1} = X'^{\leqslant}_{i-1}.$$

For all $P \subseteq C$, $y \in U$, this implies:

$$\frac{|D^{+}_{P}(y) \cap X^{\leqslant}_{i-1}|}{|X^{\leqslant}_{i-1}|} = \frac{|D'^{+}_{P}(y) \cap X'^{\leqslant}_{i-1}|}{|X'^{\leqslant}_{i-1}|} \iff \epsilon^{P}_{X^{\geqslant}_i}(y) = \epsilon^{P}_{X'^{\geqslant}_i}(y),$$

where $D'^{+}_{P}(y)$ denotes $P$-positive dominance cone of object $y$ when considering universe $U \cup X^{\Delta\geqslant}_i$. $\square$

**Proof of Theorem 5.3.** Analogous to Proof of Theorem 5.2 which is carried out for sets of objects $X^{\geqslant}_i$ and $X'^{\geqslant}_i$. In this case, sets of objects $X^{\geqslant}_{i+1}$ and $X'^{\geqslant}_{i+1}$ are considered instead of sets $X^{\leqslant}_{i-1}$ and $X'^{\leqslant}_{i-1}$, respectively. $\square$

**Proof of Theorem 5.4.** Let us consider $y_1, y_2 \in U$ such that $y_1 D_P y_2$, $P \subseteq C$. From the definition of dominance cone $D_P^+(y)$, $y \in U$,

$$D_P^+(y_1) \subseteq D_P^+(y_2).$$

For $X_i^{\geqslant}, X_{i-1}^{\leqslant} \subseteq U$, this implies:

$$D_P^+(y_1) \cap X_{i-1}^{\leqslant} \subseteq D_P^+(y_2) \cap X_{i-1}^{\leqslant} \Rightarrow |D_P^+(y_1) \cap X_{i-1}^{\leqslant}| \leqslant |D_P^+(y_2) \cap X_{i-1}^{\leqslant}|$$
$$\Rightarrow \frac{|D_P^+(y_1) \cap X_{i-1}^{\leqslant}|}{|X_{i-1}^{\leqslant}|} \leqslant \frac{|D_P^+(y_2) \cap X_{i-1}^{\leqslant}|}{|X_{i-1}^{\leqslant}|} \iff \epsilon_{X_i^{\geqslant}}^{P}(y_1) \leqslant \epsilon_{X_i^{\geqslant}}^{P}(y_2).$$

The proof for downward union $X_i^{\leqslant}$ is analogical, but starts from the observation that for negative dominance cone $D_P^-(y)$, $y \in U$,

$$D_P^-(y_1) \supseteq D_P^-(y_2). \qquad \square$$

**Proof of Theorem 5.5.** As it was already proved in Proof of Theorem 5.1, for all $P \subseteq P' \subseteq C$, $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, $y \in U$,

$$\epsilon_{X_i^{\geqslant}}^{P}(y) \geqslant \epsilon_{X_i^{\geqslant}}^{P'}(y)$$

and

$$\epsilon_{X_i^{\leqslant}}^{P}(y) \geqslant \epsilon_{X_i^{\leqslant}}^{P'}(y).$$

Thus,

$$\epsilon_{X_i^{\geqslant}}^{*P}(y) = \max_{j \leqslant i} \epsilon_{X_j^{\geqslant}}^{P}(y) \geqslant \max_{j \leqslant i} \epsilon_{X_j^{\geqslant}}^{P'}(y) = \epsilon_{X_i^{\geqslant}}^{*P'}(y)$$

and

$$\epsilon_{X_i^{\leqslant}}^{*P}(y) = \max_{j \geqslant i} \epsilon_{X_j^{\leqslant}}^{P}(y) \geqslant \max_{j \geqslant i} \epsilon_{X_j^{\leqslant}}^{P'}(y) = \epsilon_{X_i^{\leqslant}}^{*P'}(y). \qquad \square$$

**Proof of Theorem 5.6.** New objects are introduced to union of classes $X_i^{\geqslant} \subseteq U$. Thus, for all sets of objects $X_i^{\geqslant} \subseteq U$, $X_i'^{\geqslant} = X_i^{\geqslant} \cup X_i^{\Delta \geqslant}$, where $X_i^{\Delta \geqslant} \cap U = \emptyset$,

$$X_{i-1}^{\leqslant} = X_{i-1}'^{\leqslant}.$$

For all $P \subseteq C$, $y \in U$, this implies:

$$\epsilon_{X_i^{\geqslant}}^{*P}(y) = \max_{j \leqslant i} \frac{|D_P^+(y) \cap X_{j-1}^{\leqslant}|}{|X_{j-1}^{\leqslant}|} = \max_{j \leqslant i} \frac{|D_P'^+(y) \cap X_{j-1}'^{\leqslant}|}{|X_{j-1}'^{\leqslant}|} = \epsilon_{X_i'^{\geqslant}}^{*P}(y),$$

where $D_P'^+(y)$ denotes $P$-positive dominance cone of object $y$ when considering universe $U \cup X_i^{\Delta \geqslant}$. $\square$

**Proof of Theorem 5.7.** Analogous to Proof of Theorem 5.6 which is carried out for sets of objects $X_i^{\geqslant}$ and $X_i'^{\geqslant}$. In this case, sets of objects $X_{i+1}^{\geqslant}$ and $X_{i+1}'^{\geqslant}$ are considered instead of sets $X_{i-1}^{\leqslant}$ and $X_{i-1}'^{\leqslant}$, respectively. $\square$

**Proof of Theorem 5.8.** Let us consider $P \subseteq C$, $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, $j \leqslant i$, $y \in U$. Since $j \leqslant i$,

$$\epsilon_{X_i^{\geqslant}}^{*P}(y) = \max_{k \leqslant i} \frac{|D_P^+(y) \cap X_{k-1}^{\leqslant}|}{|X_{k-1}^{\leqslant}|} \geqslant \max_{k \leqslant j} \frac{|D_P^+(y) \cap X_{k-1}^{\leqslant}|}{|X_{k-1}^{\leqslant}|} = \epsilon_{X_j^{\geqslant}}^{*P}(y). \qquad \square$$

**Proof of Theorem 5.9.** Analogous to Proof of Theorem 5.8. $\square$

**Proof of Theorem 5.10.** Let us consider $y_1, y_2 \in U$ such that $y_1 D_P y_2$, $P \subseteq C$. From the definition of dominance cone $D_P^+(y)$, $y \in U$,

$$D_P^+(y_1) \subseteq D_P^+(y_2).$$

For $X_i^{\geqslant}, X_{i-1}^{\leqslant} \subseteq U$, this implies:

$$D_P^+(y_1) \cap X_{i-1}^{\leqslant} \subseteq D_P^+(y_2) \cap X_{i-1}^{\leqslant} \Rightarrow \frac{|D_P^+(y_1) \cap X_{i-1}^{\leqslant}|}{|X_{i-1}^{\leqslant}|} \leqslant \frac{|D_P^+(y_2) \cap X_{i-1}^{\leqslant}|}{|X_{i-1}^{\leqslant}|}$$
$$\Rightarrow \max_{k \leqslant i} \frac{|D_P^+(y_1) \cap X_{k-1}^{\leqslant}|}{|X_{k-1}^{\leqslant}|} \leqslant \max_{k \leqslant i} \frac{|D_P^+(y_2) \cap X_{k-1}^{\leqslant}|}{|X_{k-1}^{\leqslant}|} \iff \epsilon_{X_i^{\geqslant}}^{*P}(y_1) \leqslant \epsilon_{X_i^{\geqslant}}^{*P}(y_2).$$

The proof for downward union $X_i^{\leqslant}$ is analogical, but starts from the observation that for negative dominance cone $D_P^-(y)$, $y \in U$,

$$D_P^-(y_1) \supseteq D_P^-(y_2). \qquad \square$$

**Proof of Theorem 5.11.** For each object $y \in X_i^{\geqslant}$, $D_P^+(y) \subseteq X_i^{\geqslant}$ iff $\epsilon_{X_i^{\geqslant}}^{*P}(y) = 0$. For each object $y \in X_i^{\leqslant}$, $D_P^-(y) \subseteq X_i^{\leqslant}$ iff $\epsilon_{X_i^{\leqslant}}^{*P}(y) = 0$. □

**Proof of Theorem 5.12.** Analogous to Proof of Theorem 5.1 for measures $\epsilon_{X_i^{\geqslant}}^{P}(y)$ and $\epsilon_{X_i^{\leqslant}}^{P}(y)$ – only the common denominators in fractions are changed from $|X_{i-1}^{\leqslant}|$ and $|X_{i+1}^{\geqslant}|$ to $|X_i^{\geqslant}|$ and $|X_i^{\leqslant}|$, respectively. □

**Proof of Theorem 5.13.** New objects are introduced to union of classes $X_i^{\geqslant} \subseteq U$. Thus, for all sets of objects $X_i^{\geqslant} \subseteq U$, $X_i'^{\geqslant} = X_i^{\geqslant} \cup X_i^{\Delta\geqslant}$, where $X_i^{\Delta\geqslant} \cap U = \emptyset$,

$$X_{i-1}^{\leqslant} = X_{i-1}'^{\leqslant}, \quad |X_i^{\geqslant}| < |X_i'^{\geqslant}|.$$

This implies that for all $P \subseteq C$, $y \in U$:

$$\frac{|D_P^+(y) \cap X_{i-1}^{\leqslant}|}{|X_i^{\geqslant}|} > \frac{|D_P'^+(y) \cap X_{i-1}'^{\leqslant}|}{|X_i'^{\geqslant}|} \iff \epsilon_{X_i^{\geqslant}}'^{P}(y) > \epsilon_{X_i'^{\geqslant}}'^{P}(y),$$

where $D_P'^+(y)$ denotes $P$-positive dominance cone of object $y$ when considering universe $U \cup X_i^{\Delta\geqslant}$. □

**Proof of Theorem 5.14.** Analogous to Proof of Theorem 5.13, carried out for sets of objects $X_i^{\geqslant}, X_i'^{\geqslant}$. Here, sets of objects $X_{i+1}^{\geqslant}$, $X_{i+1}'^{\geqslant}$ and cardinalities of sets $|X_i^{\leqslant}|$, $|X_i'^{\leqslant}|$ are taken into account instead of sets $X_{i-1}^{\leqslant}$, $X_{i-1}'^{\leqslant}$ and cardinalities $|X_i^{\geqslant}|$, $|X_i'^{\geqslant}|$, respectively. □

**Proof of Theorem 5.15.** Let us consider $P \subseteq C$, $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, $j \leqslant i$, $y \in U$. Since $j \leqslant i$,

$$X_i^{\geqslant} \subseteq X_j^{\geqslant} \quad \text{and} \quad X_{i-1}^{\leqslant} \supseteq X_{j-1}^{\leqslant}.$$

This implies:

$$\frac{|D_P^+(y) \cap X_{i-1}^{\leqslant}|}{|X_i^{\geqslant}|} \geqslant \frac{|D_P^+(y) \cap X_{j-1}^{\leqslant}|}{|X_j^{\geqslant}|} \iff \epsilon_{X_i^{\geqslant}}'^{P}(y) \geqslant \epsilon_{X_j^{\geqslant}}'^{P}(y). \qquad \square$$

**Proof of Theorem 5.16.** Analogous to Proof of Theorem 5.15. □

**Proof of Theorem 5.17.** Analogous to Proof of Theorem 5.4 for measures $\epsilon_{X_i^{\geqslant}}^{P}(y)$ and $\epsilon_{X_i^{\leqslant}}^{P}(y)$ – only the common denominators in fractions are changed from $|X_{i-1}^{\leqslant}|$ and $|X_{i+1}^{\geqslant}|$ to $|X_i^{\geqslant}|$ and $|X_i^{\leqslant}|$, respectively. □

**Proof of Theorem 5.18.** For each object $y \in X_i^{\geqslant}$, $D_P^+(y) \subseteq X_i^{\geqslant}$ iff $\epsilon_{X_i^{\geqslant}}'^{P}(y) = 0$. For each object $y \in X_i^{\leqslant}$, $D_P^-(y) \subseteq X_i^{\leqslant}$ iff $\epsilon_{X_i^{\leqslant}}'^{P}(y) = 0$. □

**Proof of Theorem 5.19.** For all $P \subseteq P' \subseteq C$, $X_i^{\geqslant} \subseteq U$, $y \in U$,

$$\overline{\mu}_{X_i^{\geqslant}}^{P}(y) = \max_{\substack{R \subseteq P, \\ z \in D_R^-(y) \cap X_i^{\geqslant}}} \frac{|D_R^+(z) \cap X_i^{\geqslant}|}{|D_R^+(z)|} \leqslant \max_{\substack{R \subseteq P' \\ z \in D_R^-(y) \cap X_i^{\geqslant}}} \frac{|D_R^+(z) \cap X_i^{\geqslant}|}{|D_R^+(z)|} = \overline{\mu}_{X_i^{\geqslant}}^{P'}(y).$$

The same can be proved for measure $\overline{\mu}_{X_i^{\leqslant}}^{P}(y)$. □

**Proof of Theorem 5.20.** Analogous to Proof of Theorem 4.8. Let us consider $P \subseteq C$, $X_i^{\geqslant} \subseteq U$, $X_i'^{\geqslant} = X_i^{\geqslant} \cup X_i^{\Delta\geqslant}$, $X_i^{\Delta\geqslant} \cap U = \emptyset$, $y \in U$. Since all new objects are added to union of classes $X_i^{\geqslant}$, both numerator and denominator of fraction

$$\frac{\left|D_P^+(y) \cap X_i^{\geqslant}\right|}{\left|D_P^+(y)\right|} = \mu_{X_i^{\geqslant}}^{P}(y)$$

can increase only with the same number $k \geqslant 0$, equal to difference $|D_P'^+(y)| - |D_P^+(y)|$:

$$\frac{\left|D_P^+(y) \cap X_i^{\geqslant}\right| + k}{\left|D_P^+(y)\right| + k} = \frac{\left|D_P'^+(y) \cap X_i'^{\geqslant}\right|}{\left|D_P'^+(y)\right|} = \mu_{X_i'^{\geqslant}}^{P}(y),$$

where $D_P'^+(y)$ denotes the set of objects dominating object $y$ when considering set of attributes $P$ and universe $U \cup X_i^{\Delta\geqslant}$. Using the same reasoning as in Proof of Theorem 4.8, we can show that

$$\mu_{X_i^{\geqslant}}^{P}(y) \leqslant \mu_{X_i'^{\geqslant}}^{P}(y). \tag{36}$$

Thus,

$$\overline{\mu}_{X_i^{\geqslant}}^{P}(y) = \max_{\substack{R \subseteq P, \\ z \in D_R^-(y) \cap X_i^{\geqslant}}} \mu_{X_i^{\geqslant}}^{R}(z) \leqslant \max_{\substack{R \subseteq P, \\ z \in D_R'^-(y) \cap X_i'^{\geqslant}}} \mu_{X_i'^{\geqslant}}^{R}(z) = \overline{\mu}_{X_i'^{\geqslant}}^{P}(y). \qquad \square$$

**Proof of Theorem 5.21.** Analogous to Proof of Theorem 5.20 – only the upward unions are changed to downward unions and positive dominance cones are changed to negative dominance cones, respectively. □

**Proof of Theorem 5.22.** Let us consider $P \subseteq C$, $X_i^{\geqslant} \subseteq X_j^{\geqslant} \subseteq U$, $j \leqslant i$, $y \in U$. Since $X_i^{\geqslant} \subseteq X_j^{\geqslant}$,

$$\overline{\mu}^P_{X_i^{\geqslant}}(y) = \max_{\substack{R \subseteq P,\\ z \in D_R^-(y) \cap X_i^{\geqslant}}} \frac{|D_R^+(z) \cap X_i^{\geqslant}|}{|D_R^+(z)|} \leqslant \max_{\substack{R \subseteq P,\\ z \in D_R^-(y) \cap X_j^{\geqslant}}} \frac{|D_R^+(z) \cap X_j^{\geqslant}|}{|D_R^+(z)|} = \overline{\mu}^P_{X_j^{\geqslant}}(y). \qquad \square$$

**Proof of Theorem 5.23.** Analogous to Proof of Theorem 5.22. Unions of classes $X_i^{\leqslant} \subseteq X_j^{\leqslant} \subseteq U$ are considered. □

**Proof of Theorem 5.24.** Let us consider $y_1, y_2 \in U$ such that $y_1 D_P y_2$, $P \subseteq C$. From the definitions of dominance cones $D_P^+(y)$ and $D_P^-(y)$, $y \in U$,

$$D_P^+(y_1) \subseteq D_P^+(y_2) \text{ and } D_P^-(y_1) \supseteq D_P^-(y_2).$$

For $X_i^{\geqslant}, X_i^{\leqslant} \subseteq U$, this implies:

$$\forall R \subseteq P : D_R^-(y_1) \supseteq D_R^-(y_2) \Rightarrow \forall R \subseteq P : D_R^-(y_1) \cap X_i^{\geqslant} \supseteq D_R^-(y_2) \cap X_i^{\geqslant} \Rightarrow \{(R,z) : R \subseteq P, z \in D_R^-(y_1) \cap X_i^{\geqslant}\} \supseteq \{(R,z) : R \subseteq P, z \in D_R^-(y_2) \cap X_i^{\geqslant}\} \Rightarrow \max_{\substack{R \subseteq P,\\ z \in D_R^-(y_1) \cap X_i^{\geqslant}}} \frac{|D_R^+(z) \cap X_i^{\geqslant}|}{|D_R^+(z)|} \geqslant \max_{\substack{R \subseteq P,\\ z \in D_R^-(y_2) \cap X_i^{\geqslant}}} \frac{|D_R^+(z) \cap X_i^{\geqslant}|}{|D_R^+(z)|} \Longleftrightarrow \overline{\mu}^P_{X_i^{\geqslant}}(y_1) \geqslant \overline{\mu}^P_{X_i^{\geqslant}}(y_2),$$

$$\forall R \subseteq P : D_R^+(y_1) \subseteq D_R^+(y_2) \Rightarrow \forall R \subseteq P : D_R^+(y_1) \cap X_i^{\leqslant} \subseteq D_R^+(y_2) \cap X_i^{\leqslant} \Rightarrow \{(R,z) : R \subseteq P, z \in D_R^+(y_1) \cap X_i^{\leqslant}\} \subseteq \{(R,z) : R \subseteq P, z \in D_R^+(y_2) \cap X_i^{\leqslant}\} \Rightarrow \max_{\substack{R \subseteq P,\\ z \in D_R^+(y_1) \cap X_i^{\leqslant}}} \frac{|D_R^-(z) \cap X_i^{\leqslant}|}{|D_R^-(z)|} \leqslant \max_{\substack{R \subseteq P,\\ z \in D_R^+(y_2) \cap X_i^{\leqslant}}} \frac{|D_R^-(z) \cap X_i^{\leqslant}|}{|D_R^-(z)|} \Longleftrightarrow \overline{\mu}^P_{X_i^{\leqslant}}(y_1) \leqslant \overline{\mu}^P_{X_i^{\leqslant}}(y_2). \qquad \square$$

**Proof of Theorem 5.25.** For each object $y \in X_i^{\geqslant}$, $D_P^+(y) \subseteq X_i^{\geqslant}$ iff $\overline{\mu}^P_{X_i^{\geqslant}}(y) = 1$. For each object $y \in X_i^{\leqslant}$, $D_P^-(y) \subseteq X_i^{\leqslant}$ iff $\overline{\mu}^P_{X_i^{\leqslant}}(y) = 1$. □

## References

[1] J. Błaszczyński, S. Greco, R. Słowiński, M. Szeląg, On variable consistency dominance-based rough set approaches, in: S. Greco, Y. Hata, S. Hirano, M. Inuiguchi, S. Miyamoto, H.S. Nguyen, R. Słowiński (Eds.), Rough Sets and Current Trends in Computing, LNAI, vol. 4259, Springer-Verlag, 2006, pp. 191–202.
[2] J. Błaszczyński, S. Greco, R. Słowiński, M. Szeląg, Monotonic variable consistency rough set approaches, in: J. Yao, P. Lingras, W. Wu, M. Szczuka, N.J. Cercone, D. Śle¸zak (Eds.), Rough Sets and Knowledge Technology, LNAI, vol. 4481, Springer-Verlag, 2007, pp. 126–133.
[3] G. Brewka, Nonmonotonic Reasoning: Logical Foundations of Commonsense, Cambridge University Press, 1991.
[4] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106 (1) (1998) 109–137.
[5] B. Fitelson, Studies in Bayesian Confirmation Theory, Ph. D. Thesis, University of Wisconsin-Madison, 2001.
[6] B. Fitelson, Likelihoodism, Bayesianism, and relational confirmation, Synthese 156 (2007) 473–489.
[7] M.L. Ginsberg, Readings in Nonmonotonic Reasoning, Morgan Kaufman, Los Altos, CA, 1987.
[8] S. Greco, B. Matarazzo, R. Słowiński, The use of rough sets and fuzzy sets in MCDM, in: T. Gal, T. Hanne, T. Stewart (Eds.), Advances in Multiple Criteria Decision Making, Kluwer Academic Publishers, Boston, 1999, pp. 14.1–14.59 (Chapter 14).
[9] S. Greco, B. Matarazzo, R. Słowiński, Rough sets theory for multicriteria decision analysis, European Journal of Operational Research 129 (1) (2001) 1–47.
[10] S. Greco, B. Matarazzo, R. Słowiński, J. Stefanowski, Variable consistency model of dominance-based rough sets approach, in: W. Ziarko, Y. Yao (Eds.), Rough Sets and Current Trends in Computing, LNAI, vol. 2005, Springler-Verlag, Berlin, 2001, pp. 170–181.
[11] S. Greco, B. Matarazzo, R. Słowiński, Rough approximation by dominance relations, International Journal of Intelligent Systems 17 (2) (2002) 153–171.
[12] S. Greco, Z. Pawlak, R. Słowiński, Can Bayesian confirmation measures be useful for rough set decision rules?, Engineering Applications of Artificial Intelligence 17 (2004) 345–361.
[13] S. Greco, R. Słowiński, Y.Y. Yao, Bayesian Decision theory for dominance-based rough set approach, in: J. Yao, P. Lingras, W. Wu, M. Szczuka, N.J. Cercone, D. Śle¸zak (Eds.), Rough Sets and Knowledge Technology, LNAI, vol. 4481, Springer-Verlag, 2007, pp. 134–141.
[14] S. Greco, B. Matarazzo, R. Słowiński, Rough membership and Bayesian confirmation measures for parameterized rough sets, International Journal of Approximate Reasoning 49 (2) (2008) 285–300.
[15] C.G. Hempel, Studies in the logic of confirmation (I), Mind 54 (1945) 1–26.
[16] Q. Hu, Y. Daren, X. Zongxia, J. Liu, Fuzzy probabilistic approximation spaces and their information measures, IEEE Transactions on Fuzzy Systems 14 (2) (2006) 191–201.
[17] M. Inuiguchi, Structure-based approaches to attribute reduction in variable precision rough set models, in: Granular Computing, 2005 IEEE International Conference, vol. 1, pp. 34–39.
[18] D. Makinson, Bridges from Classical to Nonmonotonic Logic, College Publications, 2005.
[19] Z. Pawlak, Rough sets, International Journal of Information and Computer Sciences 11 (1982) 341–356.
[20] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Norwell, MA, USA, 1991.
[21] Z. Pawlak, A. Skowron, Rough membership functions, in: R.R. Yager, M. Fedrizzi, J. Kacprzyk (Eds.), Advances in the Dempster–Shafer Theory of Evidence, Wiley, New York, 1994, pp. 251–271.
[22] Y. Qian, J. Liang, Ch. Dang, H. Zhang, M. Jianmin, On the evaluation of the decision performance of an incomplete decision table, Data and Knowledge Engineering 65 (3) (2008) 373–400.
[23] Y. Qian, J. Liang, Ch. Dang, Consistency measure, inclusion degree and fuzzy measure in decision tables, Fuzzy Sets and Systems 159 (2008) 2353–2377.

[24] R. Słowiński, S. Greco, B. Matarazzo, Rough set based decision support, in: E.K. Burke, G. Kendall (Eds.), Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques, Springer-Verlag, New York, 2005, pp. 475–527 (Chapter 16).
[25] D. Śle¸zak, Rough Sets and Bayes Factor, Transactions on Rough Sets III, LNCS, vol. 3400, Springer-Verlag, 2005. pp. 202–229.
[26] D. Śle¸zak, W. Ziarko, The investigation of the Bayesian rough set model, International Journal of Approximate Reasoning 40 (2005) 81–91.
[27] S.K.M. Wong, W. Ziarko, Comparison of the probabilistic approximate classification and the fuzzy set model, Fuzzy Sets and Systems 21 (1987) 357–362.
[28] W. Ziarko, Variable precision rough set model, Journal of Computer and System Sciences 46 (1) (1993) 39–59.
[29] W. Ziarko, Stochastic approach to rough set theory, in: S. Greco, Y. Hata, S. Hirano, M. Inuiguchi, S. Miyamoto, H.S. Nguyen, R. Słowiński (Eds.), Rough Sets and Current Trends in Computing, LNAI, vol. 4259, Springer-Verlag, 2006, pp. 38–48.
[30] Y.Y. Yao, Decision-theoretic rough set models, in: J. Yao, P. Lingras, W. Wu, M. Szczuka, N.J. Cercone, D. Śle¸zak (Eds.), Rough Sets and Knowledge Technology, LNAI, vol. 4481, Springer-Verlag, 2007, pp. 11–12.
[31] Y.Y. Yao, Probabilistic rough set approximations, International Journal of Approximate Reasoning 49 (2) (2008) 255–271.